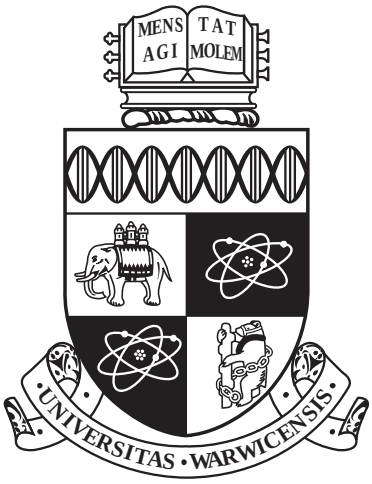

Evidence for the Higgs Boson Yukawa Coupling to Tau Leptons with the ATLAS Detector

by

Carl Thomas Jeske

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

July 2015

For Dad

Keeping a promise

Contents

List of Tables

List of Figures

Acknowledgments

I'll start by thanking the entire ATLAS group at the University of Warwick. During my PhD I've had the pleasure of witnessing the group grow from a just a few people to a dozen. I'd especially like to thank Graham, Michel and Elisabetta who I've worked closely with over the years, your help has been much appreciated. Big thanks to my supervisor, Dr. Sinead Farrington, whose guidance has been vital to my progress over the last few years. Thank you also to my second supervisor, Professor Paul Harrison.

I'd also like to thank the ATLAS $H \to \tau\tau$ analysis group. A lot of hard work over the last few years has given us a result we can all be proud of. Special thanks to Dimitris, Alex, John, Jacob and Henrik of the CN and OWLS teams, you made working on the analysis software a pleasure.

With my Dad's illness, the last few years have been especially difficult for my family. I wish I'd been able to spend more time with them and I'm sorry you're not here to see this finished Dad. Thank you to friends and family, you've been so helpful and supportive in this difficult time.

Thank you also to all my friends, both new and old. Living in Switzerland and France was a real adventure and the friends I made there made it all the better. Special thanks to Rich, Shaun, Gary and Sam.

Finally, I'd like to thank Tamsin. My life changed when I met you and being with you makes me so happy. You've been incredibly supportive over the last year and kept my spirits up when things got tough.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

The work of the LHC and ATLAS are highly collaborative projects. I have contributed to the healthy running of the ATLAS experiment through my muon trigger service task and taking shifts in the ATLAS control room. Figures 4.2 and 4.3 and the results and conclusions of Section 4.3 are entirely my own work.

The $H \to \tau\tau$ analysis is also the result of a collaboration between a few dozen members of ATLAS. In the modern world of particle physics, it is infeasible to perform such a complex analysis individually. My primary contribution was the development and maintenance of the software used to produce the data ntuples used in the $H \to \tau_\ell \tau_{had}$ analysis.

This development work involved work on implementing almost all major components of the analysis. This includes the selection, categorisation, background estimation and the appropriate treatment of all experimental systematic uncertainties. I was also responsible for running the software to produce the ntuples when noteworthy updates were made.

Other tasks included implementing the latest recommendations on how to treat analysis objects and any updates to the $H \to \tau_\ell \tau_{had}$ analysis procedure, as well as providing technical expertise when unique samples were required for specialist studies. An example of such a specialist study is my work adapting the ntuple code to run on the MC embedding samples. This was an important step in the full validation of the embedding procedure.

One significant aspect of this work was the incorporation of the 7 TeV data into the analysis. This required a large amount of work ensuring that all the selection and correction procedures were fully up to date. Without this work it would have been impossible to include the 7 TeV data in the recently published paper.

I also contributed to the calculations of the theoretical uncertainties, discussed in Section 8.2.4. I personally calculated the uncertainties relating to the choice of underlying event, parton showering, PDF and MC generator. I also produced validation plots for many variables in a variety of analysis categories. This was important to ensure that the data was well modelled by the background predictions.

Abstract

Results of a search for $H \to \tau_\ell \tau_{had}$ are presented, using the full Run 1 ATLAS dataset of pp collisions. This dataset consists of an integrated luminosity of 4.5 fb^{-1} at $\sqrt{s} = 7$ TeV and 20.3 fb^{-1} at $\sqrt{s} = 8$ TeV. For a Higgs mass of 125.36 GeV, a 2.3σ (2.3σ) excess over the predicted background is observed (expected). This corresponds to a signal strength parameter of $\mu = 1.0 \pm 0.5$. When the result is combined with complementary searches for $H \to \tau_\ell \tau_\ell$ and $H \to \tau_{had}\tau_{had}$, the observed (expected) excess is 4.5σ (3.4σ). For the combined analysis, μ is measured to be 1.4 ± 0.4. This result constitutes the first direct evidence of Yukawa couplings and is fully compatible with the Standard Model expectation.

Chapter 1

Theoretical Background

1.1 Introduction

The current pinnacle of elementary particle physics theory is known as the Standard Model of Particle Physics (SM). It describes the behaviour of three of the four fundamental interactions and has been experimentally verified to a very high precision. The electromagnetic, weak and strong interactions are the three interactions described by the Standard Model, while gravitational interactions are omitted.

The Standard Model includes a number of elementary particles which can be divided into two distinct subgroups by their spin. In this context spin refers to the quantised angular momentum inherent to these fundamental particles. The spin of a particle is quantised such that it may take half-integer as well as integer values. The first subset of fundamental particles are those where the spin takes a half-integer value and these are named fermions. The second subset are referred to as bosons and these have integer values for their spins.

1.2 Fermions

As well as having spins of half integer values, other characteristics of fermions include that their behaviour is described by Fermi-Dirac statistics and that they obey the Pauli exclusion principle. This discussion focuses on fundamental fermionic particles however it should be noted that composite particles can also be fermions. Fundamental fermionic particles themselves can be split into two subgroups, the leptons and the quarks.

1.2.1 Leptons

The lowest mass, charged lepton is the electron. It has a mass of 0.511 MeV [1] and an electric charge of -1.6×10^{-19} Coulombs [1]. Expressed in the units of the charge of the proton, e, as is standard in particle physics and throughout this thesis, the electron charge is simply -1. On a similar note, natural units are also used throughout this work, meaning $\hbar = c = 1$ [2]. Like all fundamental particles, the electron is assumed to be point-like. It interacts with other matter through the electromagnetic and weak interactions although, as the lightest charged lepton, the electron is a stable particle.

As well as the electron there are a further two generations, or flavours, of charged leptons. Each generation has similar characteristics with the main difference being the mass, which increases with each generation. The second generation charged lepton is called the muon and has a mass of 105.7 MeV [1]. A further way in which the heavier charged leptons differ from the electron is that they are unstable particles and are able to decay to electrons. Throughout this thesis, the electron and muon will be referred to as light leptons, denoted with the symbol ℓ. The final and most massive charged lepton is known as the tau or the tauon. This particle is also unstable, being able to decay to an electron or muon, and has a mass of $1,776.8$ MeV [1].

Alongside the charged leptons, there are also three neutral leptons named neutrinos, one corresponding to each flavour of lepton. Long thought to be massless, the existence of neutrino oscillations has demonstrated that neutrinos, while very low mass, are not entirely massless. Neutrino oscillation is the phenomenon in which neutrinos are able to change flavour between creation and detection [3]. For example electron neutrinos produced in the Sun may be detected on Earth as muon neutrinos. As neutral particles, neutrinos only interact with other matter through the weak interaction.

Each generation of leptons has an associated lepton number. For example, electrons and electron neutrinos have an electron number equal to $+1$ while the other leptons have an electron number of 0. The conservation of the three lepton numbers can only be violated by neutrino oscillation in the Standard Model. Total lepton number (the sum of electron number, muon number and tau number) must be conserved.

Each particle listed above also has a corresponding antiparticle. Antiparticles are identical to their particle counterparts except with a number of physical quantities reversed. For instance the positron, the electron's antiparticle, has the same mass as the electron but a charge of $+1$ and an electron number of -1. The

Particle	Symbol	Electric Charge	Mass	Gen.
Electron	$e^{\pm}$	± 1	0.511 MeV	1st
Muon	$\mu^{\pm}$	± 1	105.658 MeV	2nd
Tau	$\tau^{\pm}$	± 1	1,776.82 MeV	3rd
Electron neutrino	ν_e	0	< 2 eV	1st
Muon neutrino	ν_μ	0	< 2 eV	2nd
Tau neutrino	ν_τ	0	< 2 eV	3rd

Table 1.1: The leptons of the Standard Model [1].

Particle	Symbol	Electric Charge	Mass	Type	Gen.
Up	u	$+\frac{2}{3}$	$2.3^{+0.7}_{-0.5}$ MeV	up-type	1st
Charm	c	$+\frac{2}{3}$	1.275 ± 0.025 GeV	up-type	2nd
Top	t	$+\frac{2}{3}$	$173.21 \pm 0.51 \pm 0.71$ GeV	up-type	3rd
Down	d	$-\frac{1}{3}$	$4.8^{+0.5}_{-0.3}$ MeV	down-type	1st
Strange	s	$-\frac{1}{3}$	95 ± 5 MeV	down-type	2nd
Bottom	b	$-\frac{1}{3}$	4.18 ± 0.03 GeV	down-type	3rd

Table 1.2: The quarks of the Standard Model [1].

properties of the leptons of the Standard Model are summarised in Table 1.1.

1.2.2 Quarks

Similar to the leptons, the quarks come in three generations with two particles in each generation. One particle in each generation has a charge of $+\frac{2}{3}$ and is known as an up-type quark, while the other is a down-type quark with a charge of $-\frac{1}{3}$. Each quark also has an associated antiparticle with the inverse charge. As quarks carry a colour charge they can interact via the strong force as well as the electromagnetic and weak forces.

Due to colour confinement, quarks are only observed in colour neutral bound states called hadrons. Three quarks (one red, one green and one blue) can join to form a baryon, with the most stable baryons being protons and neutrons. A quark with a colour can also bind with an antiquark of the corresponding anticolour to form a bound state termed a meson. The properties of the quarks are summarised in Table 1.2.

All stable matter in the known universe is made up of first generation fermions. Protons and neutrons are made of up and down quarks only and these combine with

electrons to form atoms. The reason behind this, if one even exists, is not well understood.

1.3 Bosons

While fermions make up the matter of the Universe, bosons are responsible for carrying the three forces of the Standard Model. Each of the three fundamental interactions are mediated by different bosons. Particles interact by exchanging bosons, with the bosons often being virtual particles. Bosons are subject to Bose-Einstein statistics.

Electromagnetic interactions are mediated by the photon. The photon has a spin of 1 and is a neutral particle. Despite carrying no charge itself, the photon couples to the electric charge of elementary particles. As photons are massless, the electromagnetic force has an infinite range.

The weak force has two spin-1 force carriers. The Z^0 is neutral while the $W^{\pm}$ has a charge of ± 1. The $W^{\pm}$ is especially interesting as fermions coupling to it may change flavour. For example, a muon may decay to a $W^{\pm}$ and a muon neutrino with the $W^{\pm}$ then decaying to an electron and electron antineutrino or two quarks (one up-type and one down-type).

The $W^{\pm}$ and Z^0 are both massive particles, setting them apart from the other force carriers. The mechanism by which these bosons gain their mass is called the Higgs mechanism and is discussed in more detail below.

Gluons are the force carrying bosons which mediate strong interactions. Gluons are also spin-1 particles. Unlike the photon, gluons carry a charge so are directly involved in strong interactions and can couple to themselves. These self-interactions cause a phenomenon known as confinement and mean gluon-mediated interactions can only occur over very short distances of around 10^{-15} m.

The three colour and three anticolour charges of the strong force give nine independent colour states for gluons. However, the neutral colour singlet state $(r\bar{r} + b\bar{b} + g\bar{g})/\sqrt{3}$ is not observed in nature. This leaves eight gluons, each one a different combination of the colour and anticolour charges. One commonly used representation of the eight colour states is

$$\begin{aligned}
&(r\bar{b} + b\bar{r})/\sqrt{2}, \quad -i(r\bar{b} - b\bar{r})/\sqrt{2}, \\
&(r\bar{g} + g\bar{r})/\sqrt{2}, \quad -i(r\bar{g} - g\bar{r})/\sqrt{2}, \\
&(b\bar{g} + g\bar{b})/\sqrt{2}, \quad -i(b\bar{g} - g\bar{b})/\sqrt{2}, \\
&(r\bar{r} - b\bar{b})/\sqrt{2}, \quad (r\bar{r} + b\bar{b} - 2g\bar{g})/\sqrt{6}.
\end{aligned}$$

Particle	Symbol	Electric Charge	Spin	Mass
Photon	γ	0	1	0 GeV
$W^{\pm}$	$W^{\pm}$	± 1	1	80.385 ± 0.015 GeV
Z^0	Z^0	0	1	91.1876 ± 0.0021 GeV
Gluon	g	0	1	0 GeV
Higgs	H	0	0	125.7 ± 0.4 GeV

Table 1.3: The bosons of the Standard Model [1].

The final boson included in the Standard Model is the Higgs boson. This boson's existence, although theorised in 1964 [4, 5, 6], was proved in 2012 by the ATLAS and CMS collaborations at the Large Hadron Collider (LHC) [7, 8]. The Higgs is a neutral fundamental particle and has a mass of around 125 GeV [9, 10]. There is strong evidence that the Higgs boson is a spin-0 particle [11, 12]. The Higgs boson is a key component of the Higgs mechanism, the mechanism which gives mass to the $W^{\pm}$ and Z^0 bosons. This mechanism is explained in detail below.

The properties of the bosons are summarised in Table 1.3.

1.4 Gauge Theories

In particle physics, particles and their interactions can be described using a formalism named relativistic field theory. Mathematically, systems are described by Lagrangian densities (shortened to simply Lagrangian in this thesis), denoted by the symbol $\mathscr{L}$. For example, the Lagrangian which describes a spin-$\frac{1}{2}$ field, ψ, is [13]

$$\mathscr{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi. \tag{1.1}$$

This is known as the Dirac Lagrangian.

1.4.1 Local Gauge Invariance

By inspection we can determine that Equation 1.1 is invariant under a global gauge transformation [13]

$$\psi \rightarrow e^{-iq\lambda}\psi. \tag{1.2}$$

This is known as the $U(1)$ gauge transformation. In group theory, $U(n)$ refers to the group of $n \times n$ matrices which are unitary, i.e. $U^{\dagger}U = 1$. For the one dimensional case this group can simply be expressed in the form $e^{-i\theta}$, as in the equation above.

If we however consider the case where λ depends on position, x, known as a

local gauge transformation, this invariance is broken. Instead we find [13]

$$\mathscr{L} \to \mathscr{L} + (q\bar{\psi}\gamma^{\mu}\psi)\partial_{\mu}\lambda. \tag{1.3}$$

Now, consider the Lagrangian

$$\mathscr{L} = -(q\bar{\psi}\gamma^{\mu}\psi)A_{\mu}. \tag{1.4}$$

Here A_{μ} represents a vector field which, under the local gauge transformation, transforms like

$$A_{\mu} \to A_{\mu} + \partial_{\mu}\lambda. \tag{1.5}$$

This new Lagrangian transforms like

$$\mathscr{L} \to \mathscr{L} - (q\bar{\psi}\gamma^{\mu}\psi)\partial_{\mu}\lambda \tag{1.6}$$

and, while not locally gauge invariant itself, would perfectly cancel the unwanted term in Equation 1.3. Therefore, by adding the Lagrangians of Equations 1.1 and 1.4 we can construct a locally gauge invariant Lagrangian [13]

$$\mathscr{L} = (i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi) - (q\bar{\psi}\gamma^{\mu}\psi)A_{\mu}. \tag{1.7}$$

One final step is to add the free Lagrangian for the field A_{μ}. As a vector field, the free Lagrangian takes the form [13]

$$\mathscr{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m^2A^{\nu}A_{\nu}, \tag{1.8}$$

where $F^{\mu\nu}$ is shorthand for the gauge invariant expression $\partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu}$. The second term however is not locally gauge invariant meaning that the vector field must be massless [13]. This gives a final Lagrangian of

$$\mathscr{L} = (i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi) - (q\bar{\psi}\gamma^{\mu}\psi)A_{\mu} - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{1.9}$$

So, in summary, by enforcing local $U(1)$ gauge invariance on the Dirac Lagrangian we find this requires the presence of a massless vector field. This is in fact a simple version of the Lagrangian for Quantum Electrodynamics (QED). We can see spin-$\frac{1}{2}$ particles (electrons and positrons) interacting with a vector field (photons). Locally gauge invariant theories are also known as gauge theories.

Before continuing it should be noted that Equation 1.1 can be changed to

Equation 1.7 simply by substituting the ∂_μ with the expression

$$\mathscr{D}_\mu = \partial_\mu + iqA_\mu. \tag{1.10}$$

$\mathscr{D}_\mu$ is called the covariant derivative and will be of use in later sections.

1.4.2 Expanding Local Gauge Invariance to SU(2)

In section 1.4.1 we enforced $U(1)$ local gauge invariance. With fairly minimal changes this method can be extended to the $SU(2)$ gauge transformation. This method was pioneered in 1954 by Yang and Mills [14].

The starting point is again the Dirac Lagrangian but now for two spin-$\frac{1}{2}$ particles of equal mass. These can simply be represented by a column vector

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}. \tag{1.11}$$

This gives a starting Lagrangian identical to Equation 1.1, the only difference being ψ and $\bar{\psi}$ now have two elements.

Similar to $U(1)$, the $SU(2)$ group refers to 2×2 matrices which are unitary. The addition of the S simply means that as well as the requirement of unitarity, each matrix must also have a determinant of 1. The $SU(2)$ transformation can be written in the form [13]

$$\psi \to e^{-iq\boldsymbol{\tau}\cdot\boldsymbol{\lambda}}\psi, \tag{1.12}$$

where $\boldsymbol{\tau}$ and $\boldsymbol{\lambda}$ are both three vectors. Each element in the $\boldsymbol{\tau}$ vector is a Pauli matrix [15]. The three elements of $\boldsymbol{\lambda}$ are real numbers for a global gauge transformation or functions of position, x, for a local transformation. The Dirac Lagrangian is invariant under a global $SU(2)$ transformation. As before, we will proceed with the case of the local gauge transformation and adapt Equation 1.1 to be locally $SU(2)$ invariant.

At the end of Section 1.4.1 it was stated that replacing each instance of $\partial\mu$ with the covariant derivative, $\mathscr{D}_\mu$, served as a useful first step in the quest for a locally gauge invariant Lagrangian. In the $SU(2)$ case, Equation 1.10 becomes [13]

$$\mathscr{D}_\mu = \partial_\mu + iq\boldsymbol{\tau} \cdot \boldsymbol{A}_\mu. \tag{1.13}$$

In this case, $\boldsymbol{A}_\mu$ is a three vector of vector fields.

The transformation for these vector fields required to make the Lagrangian locally gauge invariant is complicated but can be sufficiently approximated for small

$|\lambda|$. In this approximation the fields transform as [13]

$$\boldsymbol{A}_\mu \to \boldsymbol{A}_\mu + \partial_\mu \boldsymbol{\lambda} + 2q(\boldsymbol{\lambda} \times \boldsymbol{A}_\mu). \tag{1.14}$$

Continuing to follow the process in the previous section, the final component required is the addition of free field terms for the new vector fields. These take a similar form to Equation 1.8 with the difference being $F^{\mu\nu}$ must be redefined as

$$\boldsymbol{F^{\mu\nu}} = \partial^\mu \boldsymbol{A}^\nu - \partial^\nu \boldsymbol{A}^\mu - 2q(\boldsymbol{A}^\mu \times \boldsymbol{A}^\nu). \tag{1.15}$$

Note that again mass terms cannot be added without breaking local gauge invariance [13].

Combining all of the above we are left with a final, locally $SU(2)$ gauge invariant, Lagrangian

$$\mathscr{L} = (i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi) - (q\bar{\psi}\gamma^\mu \boldsymbol{\tau}\psi) \cdot \boldsymbol{A}_\mu - \frac{1}{4}\boldsymbol{F}^{\mu\nu} \cdot \boldsymbol{F}_{\mu\nu}. \tag{1.16}$$

This Lagrangian describes pairs of spin-$\frac{1}{2}$ particles with identical masses, three massless vector fields and the interactions between the two. While interesting, this model describes no known physical process. This method, however, can be applied to the $SU(3)$ case with relevant physical results.

1.4.3 SU(3) Local Gauge Invariance

As the previous model described two spin-$\frac{1}{2}$ particles with identical masses, replacing ψ in Equation 1.1 with a three vector describes three spin-$\frac{1}{2}$ particles. This has a direct physical relevance in the case of the strong force. Each quark flavour is in fact three different spin-$\frac{1}{2}$ fields, one for each of the three colour charges

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix}. \tag{1.17}$$

In the $SU(3)$ case the local gauge transformation can be written as

$$\psi \to e^{-iq\boldsymbol{\lambda}\cdot\boldsymbol{\phi}}\psi, \tag{1.18}$$

where the Gell-Mann matrices [16], $\boldsymbol{\lambda}$, replace the Pauli matrices from the $SU(2)$ example and ϕ is the position dependent gauge transformation.

Now, following an identical methodology to the one used in Section 1.4.2, we

can obtain a Lagrangian which describes the strong interaction in full. The only difference of note is that each three vector is replaced by a corresponding vector with eight elements. This requires defining the cross product for vectors with eight elements

$$(\boldsymbol{A} \times \boldsymbol{B})_i = \sum_{j,k=1}^{8} f_{ijk} A_j B_k, \tag{1.19}$$

where f_{ijk} are the structure constants for $SU(3)$.

So, finally, we have a Lagrangian describing spin-$\frac{1}{2}$ particles with three colour states, the quarks, interacting with eight massless vector fields, the gluons. Gluon self-coupling is also present in the final term. The Lagrangian is

$$\mathscr{L} = (i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi) - (q\bar{\psi}\gamma^\mu \boldsymbol{\lambda}\psi) \cdot \boldsymbol{A}_\mu - \frac{1}{4}\boldsymbol{F}^{\mu\nu} \cdot \boldsymbol{F}_{\mu\nu} \tag{1.20}$$

and takes an identical form to Equation 1.16. This Lagrangian is invariant under any $SU(3)$ local gauge transformation [13].

With the electromagnetic and strong forces both successfully described, we now require a theory which can successfully accommodate the massive gauge bosons of the weak force.

1.4.4 Non-zero Vacuum Expectation Values

Any Lagrangian density can be divided into two components, the kinetic term and a potential term

$$\mathscr{L} = \mathscr{T} - \mathscr{U}. \tag{1.21}$$

The kinetic term ($\mathscr{T}$) typically contains derivatives of the field in question whereas the potential term ($\mathscr{U}$) has no fixed form. For a given field, ϕ, with mass, m, the potential often contains a component of the form $m^2\phi^2$, known as the mass term [13].

Taking the Klein-Gordon Lagrangian,

$$\mathscr{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}m^2\phi^2, \tag{1.22}$$

as an example, the first term is the kinetic term and the potential term is simply the mass term with no additional components. The factors of a half are merely convention which have no effect on the physics. For this Lagrangian, as well as all Lagrangians discussed previously, the potential term is minimised when the field is equal to zero. A point where the potential is minimised is called a ground state.

Now, consider the Lagrangian

$$\mathscr{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) + \frac{1}{2}\mu^2\phi^2 - \frac{1}{4}\lambda^2\phi^4. \tag{1.23}$$

In this case the potential term is

$$\mathscr{U} = -\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda^2\phi^4. \tag{1.24}$$

The ϕ^2 term is the wrong sign to be a mass term and it is not immediately obvious for which value of ϕ $\mathscr{U}$ is minimised.

Through simple calculus it can be shown $\mathscr{U}$ is minimised when

$$\phi = \pm\frac{\mu}{\lambda}. \tag{1.25}$$

This is the vacuum expectation value (VEV) of the field and is non-zero.

Now, the Lagrangian is most useful when the ground state corresponds to the field being zero (and hence a VEV of zero). By performing the substitution

$$\phi = \eta \pm \frac{\mu}{\lambda} \tag{1.26}$$

we can obtain such a Lagrangian.

Following the substitution, the Lagrangian is

$$\mathscr{L} = \frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2\eta^2 \pm \mu\lambda\eta^3 - \frac{1}{4}\lambda^2\eta^4 + \frac{1}{4}\left(\frac{\mu^2}{\lambda}\right)^2. \tag{1.27}$$

We see the mass of the field is $\sqrt{2}\mu$. While Equation 1.27 clearly differs from Equation 1.23, the underlying physics described by the model is identical [13].

1.5 Spontaneous Symmetry Breaking

While the potential component of Equation 1.23 is symmetric (under $\phi \leftrightarrow -\phi$), once transformed to Equation 1.27 this is no longer the case. This is an example of spontaneous symmetry breaking, a vital component of the Higgs mechanism.

As a more tangible example, consider a rotationally symmetric bowl full of water placed in a freezer. Once the water is frozen solid the crystals of ice will all point in the same direction. The choice of direction is entirely random, without external influence, and occurs spontaneously.

For a more complex Lagrangian, spontaneous symmetry breaking proceeds

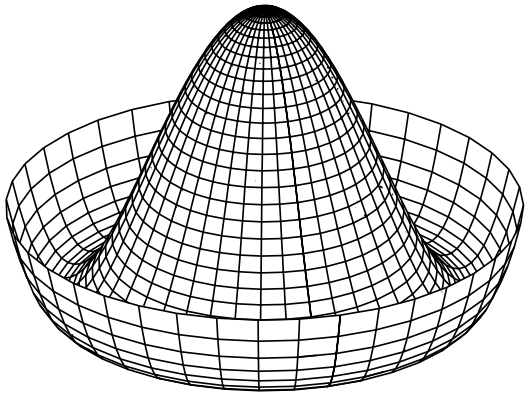

Figure 1.1: The shape of the potential for the case $\mathscr{U} = -\frac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) + \frac{1}{4}\lambda^2(\phi_1^2 + \phi_2^2)^2$. The potential is nicknamed the Mexican Hat Potential.

as follows. The Lagrangian

$$\mathscr{L} = \frac{1}{2}(\partial_\mu\phi_1)(\partial^\mu\phi_1) + \frac{1}{2}(\partial_\mu\phi_2)(\partial^\mu\phi_2) + \frac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) - \frac{1}{4}\lambda^2(\phi_1^2 + \phi_2^2)^2. \quad (1.28)$$

is similar to Equation 1.23 with the difference being it describes two fields rather than one.

The potential term takes the shape shown in Figure 1.1 and is minimised when

$$\phi_1^2 + \phi_2^2 = \frac{\mu^2}{\lambda^2}. \quad (1.29)$$

Choosing the solution of the form

$$\phi_1 = 0; \quad \phi_2 = \frac{\mu}{\lambda}; \quad (1.30)$$

we can expand about this ground state by making the substitutions

$$\phi_1 = \xi; \quad \phi_2 = \eta + \frac{\mu}{\lambda}. \quad (1.31)$$

Although the choice of solution is seemingly arbitrary, each solution describes an identical system so we are free to choose any solution. This is an example of spontaneous symmetry breaking. This particular solution was selected to simplify the algebra as we continue.

This gives a Lagrangian

$$\mathscr{L} = \left(\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \mu^2\eta^2\right) + \left(\frac{1}{2}(\partial_\mu\xi)(\partial^\mu\xi)\right)$$

$$- \left(\mu\lambda(\eta^3 + \eta\xi^2) + \frac{\lambda^2}{4}(\eta^4 + \xi^4 + 2\eta^2\xi^2) \right) + \frac{\mu^4}{4\lambda^2}. \tag{1.32}$$

It describes a system with field, η, of mass $\sqrt{2}\mu$, a massless field, ξ, and interactions between the two [13]. The massless field is a consequence of Goldstone's Theorem [17]. The theorem states that the breaking of any continuous global symmetry leads to the appearance of a massless spin-0 boson, known as a Goldstone boson. While not immediately useful in our search for a theory to describe the massive bosons of the weak interaction, this theorem forms a key part of the Higgs mechanism.

1.6 The Higgs Mechanism

Combining the principles of Section 1.4 and 1.5 will lead us to the Higgs mechanism. The two fields of Equation 1.28 can be combined into a single complex field [13]

$$\phi = \phi_1 + i\phi_2, \tag{1.33}$$

with the property

$$\phi^*\phi = \phi_1^2 + \phi_2^2. \tag{1.34}$$

Rewriting Equation 1.28 in terms of the new field we obtain the Lagrangian

$$\mathscr{L} = \frac{1}{2}(\partial_\mu\phi)^*(\partial^\mu\phi) + \frac{1}{2}\mu^2(\phi^*\phi) - \frac{1}{4}\lambda^2(\phi^*\phi)^2. \tag{1.35}$$

This Lagrangian is invariant under the global $U(1)$ gauge transformation of Equation 1.2. In fact, this transformation corresponds to a rotation about the vertical axis in Figure 1.1, an axis which the potential is clearly rotationally symmetric about.

Replacing ∂_μ with the covariant derivative, $\mathscr{D}_\mu$ (Equation 1.10), will help adapt this Lagrangian to be locally gauge invariant

$$\mathscr{L} = \frac{1}{2}\left|(\partial^\mu + iqA^\mu)\,\phi\right|^2 + \frac{1}{2}\mu^2(\phi^*\phi) - \frac{1}{4}\lambda^2(\phi^*\phi)^2 - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{1.36}$$

where $|\ \ |^2$ is shorthand for $(\ \ )^*(\ \ )$.

Expanding about a ground state, by repeating the substitution of Equation 1.31, we find

$$\begin{aligned}\mathscr{L} = &\left(\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \mu^2\eta^2\right) + \left(\frac{1}{2}(\partial_\mu\xi)(\partial^\mu\xi)\right) \\ &+ \left(-\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}q^2\left(\frac{\mu}{\lambda}\right)^2 A_\mu A^\mu\right)\end{aligned}$$

$$
\begin{aligned}
&+ \Big(q(\eta(\partial_\mu\xi) - \xi(\partial_\mu\eta))A^\mu + q^2\frac{\mu\eta}{\lambda}(A_\mu A^\mu) \\
&+ \frac{1}{2}q^2(\xi^2+\eta^2)(A_\mu A^\mu) - \lambda\mu(\eta^3+\eta\xi^2) - \frac{1}{4}\lambda^2(\eta^4+2\eta^2\xi^2+\xi^4)\Big) \\
&+ q\left(\frac{\mu}{\lambda}\right)(\partial_\mu\xi)A^\mu + \left(\frac{\mu^2}{2\lambda}\right)^2. \qquad (1.37)
\end{aligned}
$$

The first line is identical to that of Equation 1.32. It describes two spin-0 fields, the first, η, with a mass of $\sqrt{2}\mu$ and the second, ξ, massless [13]. As we saw before when enforcing local gauge invariance, a new field, A^μ, has appeared. This spin-1 field is described by the terms in the second line of the Lagrangian. Unlike the previous derivations however, this field has a non-zero mass of $q\mu/\lambda$. The third and fourth lines are interactions between the three fields.

The first term on the last line seems to describe an interaction where the ξ changes into an A and vice versa. Terms such as this tell us that the fields do not represent the fundamental particles described by the Lagrangian [13]. A real world example of this can be found in neutral Kaon oscillations. A Lagrangian formulated in terms of the K^0 and $\bar{K}^0$ has terms of this form which represent the Kaon oscillations. The true fundamental particles in this example however are the linear combinations, K_1 and K_2.

The final term in the above Lagrangian is a constant which, as usual, can be ignored [13].

While this Lagrangian has some of the desired features, others are not so nice. The massless vector field, ξ, is an unwanted component, as are the linear couplings between the ξ and A fields.

Now we have a locally gauge invariant field, it is possible to choose a gauge transformation which makes the original field, ϕ, wholly real [13]. As the imaginary component of ϕ is $\phi_2 = \xi$, this gauge transformation will remove the ξ field and its two associated issues entirely. Using the local gauge transformation

$$
\begin{aligned}
\phi &\to e^{-iq\theta}\phi \\
&\to (\cos\theta + i\sin\theta)(\phi_1 + i\phi_2) \\
&\to (\phi_1\cos\theta - \phi_2\sin\theta) + i(\phi_1\sin\theta + \phi_2\cos\theta), \qquad (1.38)
\end{aligned}
$$

to make the field real we require

$$
\phi_1\sin\theta + \phi_2\cos\theta = 0, \qquad (1.39)
$$

giving

$$\theta = -\arctan\left(\phi_2/\phi_1\right). \tag{1.40}$$

Substituting this into the Lagrangian of Equation 1.37 gives

$$\begin{aligned}\mathscr{L} = &\left(\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \mu^2\eta^2\right) + \left(-\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}q^2\left(\frac{\mu}{\lambda}\right)^2 A_\mu A^\mu\right) \\ &+ \left(q^2\frac{\mu\eta}{\lambda}(A_\mu A^\mu) + \frac{1}{2}q^2\eta^2(A_\mu A^\mu) - \lambda\mu\eta^3 - \frac{1}{4}\lambda^2\eta^4\right) \\ &+ \left(\frac{\mu^2}{2\lambda}\right)^2.\end{aligned} \tag{1.41}$$

Now, finally, we have a Lagrangian containing a massive vector field and no massless field [13]. It also predicts a massive spin-0 boson, the Higgs boson, here denoted by the symbol η. The derivation described in this section is called the Higgs mechanism and explains how bosons acquire mass in the Standard Model. The next section will discuss specifically the massive bosons which mediate the weak interaction, the $W^\pm$ and Z^0 bosons.

1.7 The Electroweak Interaction

While at low energies (below around 100 GeV) electromagnetic and weak interactions appear distinct, at higher energies the two can be treated as being two different components of the same force [13]. This combined force is called the electroweak interaction. In gauge theory, this combination is modelled as an $SU(2)\times U(1)$ group [18].

To proceed, we must apply the Higgs mechanism described above to an $SU(2) \times U(1)$ locally gauge invariant Lagrangian. This will be performed by combining the techniques used in Sections 1.4.2 and 1.6. Such a group requires ϕ to be a doublet [18]

$$\phi = \begin{pmatrix}\psi^+ \\ \psi^0\end{pmatrix}. \tag{1.42}$$

As seen above, making an $SU(2)$ Lagrangian locally gauge invariant requires the introduction of three vector fields while $U(1)$ local gauge invariance requires one vector field. Here, we will use the label $\boldsymbol{W}_\mu$ for the former and B_μ for the latter. The requirement for these fields make the covariant derivative [18]

$$\mathscr{D}_\mu = \partial_\mu + ig\boldsymbol{\tau}\cdot\boldsymbol{W}_\mu + ig'B_\mu. \tag{1.43}$$

Here, g and g' are the coupling constants of $SU(2)$ and $U(1)$ respectively.

We will continue to work with the Lagrangian shown in Equation 1.35. In this discussion, where we are looking for the theoretical mechanism which explains the masses of the electroweak bosons, the relevant terms are the ones introduced by the substitution of the covariant derivative [2]. Namely (having factored out the i present in both terms)

$$\frac{1}{2}\left|\left(g\boldsymbol{\tau}\cdot\boldsymbol{W}_\mu + g'B_\mu\right)\phi\right|^2. \tag{1.44}$$

Expanding the dot product gives

$$\frac{1}{2}\left|\begin{pmatrix} gW_\mu^3 + g'B_\mu & g(W_\mu^1 - iW_\mu^2) \\ g(W_\mu^1 + iW_\mu^2) & -gW_\mu^3 + g'B_\mu \end{pmatrix}\phi\right|^2. \tag{1.45}$$

Substituting in the value of ϕ where the potential is minimised,

$$\phi = \begin{pmatrix} 0 \\ \mu/\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{1.46}$$

and defining [2]

$$W^\pm = W^1 \mp iW^2 \tag{1.47}$$

leads to

$$\frac{1}{2}v^2\left(g^2W_\mu^+W^{-\mu} + (gW_\mu^3 - g'B_\mu)(gW^{3\mu} - g'B^\mu)\right). \tag{1.48}$$

Defining the mixing between W^3 and B as [2, 18]

$$\tan\theta_W = \frac{g'}{g} \tag{1.49}$$

and

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \tag{1.50}$$

means

$$gW^{3\mu} - g'B^\mu = \frac{g}{\cos\theta_W}Z^\mu. \tag{1.51}$$

Substituting this in gives us finally

$$\frac{1}{2}v^2\left(g^2W_\mu^+W^{-\mu} + \frac{g^2}{\cos^2\theta_W}Z_\mu Z^\mu\right). \tag{1.52}$$

We see two massive bosons, with masses

$$M_W = vg; \quad M_Z = \frac{vg}{\cos\theta_W} = \frac{M_W}{\cos\theta_W}. \tag{1.53}$$

These are the two massive bosons which mediate the weak force. Furthermore, the combination of W^3 and B which is orthogonal to Equation 1.50,

$$\sin\theta_W W_\mu^3 + \cos\theta_W B_\mu, \tag{1.54}$$

is massless and can be identified as the photon of the electromagnetic force [18].

In summary, applying the principles of local gauge invariance and expanding about a minimum in the potential to $SU(2) \times U(1)$ has led to a theory with three massive vector bosons (the Z^0, W^+ and W^-) and a massless one (the photon) [2, 18]. This is analogous to the bosons of the electroweak force. Alongside the description of the strong force in Section 1.4.3 and the theoretical origin of the Higgs in Section 1.6, we have a roughly complete description of the bosons of the SM and the origins of mass. The importance of the discovery of the Higgs boson is obvious due its key role in this theoretical framework. The final piece to be discussed is the origin of mass for the leptons.

1.8 Fermion Masses

This discussion will focus on the electron case, but may easily be expanded to all fermions. The weak interaction only applies to left-handed fermions, meaning left-handed electrons form a doublet with the electron neutrino while right-handed electrons are singlets [2]

$$\chi_L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L ; \quad \psi_R = e_R^-. \tag{1.55}$$

The fact that the Weak interaction treats left- and right-handed particles differently means a mass term of the form in the Dirac Lagrangian (Equation 1.1) will break $SU(2)$ gauge invariance. This is because the mass term would directly couple the left- and right-handed components [18].

The Lagrangian

$$\mathscr{L}_e = -G_e \left(\bar{\chi}_L \phi \psi_R + \bar{\psi}_R \phi^* \chi_L\right) \tag{1.56}$$

is $SU(2) \times U(1)$ gauge invariant (where G_e is a free constant), so can safely be added to our theory [2].

Expanding about a minimum by substituting in

$$\phi = \begin{pmatrix} 0 \\ v + H \end{pmatrix} \tag{1.57}$$

gives

$$\mathscr{L}_e = -G_e v(\bar{e}_L e_R + \bar{e}_R e_L) - G_e(\bar{e}_L e_R + \bar{e}_R e_L)H. \tag{1.58}$$

The first term is a mass term for the electron and the second is a coupling between the electron and the massive boson introduced by symmetry breaking. This demonstrates that the Higgs may explain both the masses of bosons and fermions, with the electron having mass $m_e = G_e v$ [2]. As G_e is an arbitrary constant, the theory makes no prediction as to the mass of the fermions. It does, however, suggest that fermion-Higgs couplings are proportional to the fermion masses.

The Lagrangian term discussed here is by no means necessary. It is, however, the currently preferred explanation as to the origin of the observed fermion masses. The Higgs discovery was made using its bosonic decay channels, supplemented by evidence from the fermionic decay modes. A 5σ direct observation of fermion-Higgs couplings is an important physics goal of the ATLAS and CMS experiments, key to confirming that the Higgs boson is responsible for the origin of fermion masses. Such couplings are called Yukawa couplings. $H \to \tau\tau$ is one of the most promising channels for the discovery of Yukawa couplings. This is due to the reasonable detection efficiency of ditau events, combined with their high mass when compared to the other leptons (and hence a larger Higgs coupling) [19].

1.9 Higgs Production Mechanisms at the LHC

The Higgs boson can be produced by a number of processes at a hadron collider such as the LHC. These processes are generally initiated by gluons or light quarks. Cross sections of the most common processes are summarised in Figure 1.2. The discussion here will be limited to the three processes with the largest cross sections, namely gluon-gluon fusion (ggF), vector boson fusion (VBF) and vector boson associated production (VH). The Feynman diagrams for theses processes are shown in Figure 1.3.

1.9.1 Gluon-gluon Fusion

As shown in Figure 1.3, ggF is initiated by two gluons and the Higgs is produced via a quark loop. The loop is necessary as the Higgs boson couples to mass, therefore it cannot directly couple to the massless gluons. As fermion-Higgs couplings are proportional to mass, the heaviest quarks make the largest contribution to the ggF cross section. Furthermore, a larger than expected ggF cross section can indicate new physics through the existence of heavier particles which have not been previously

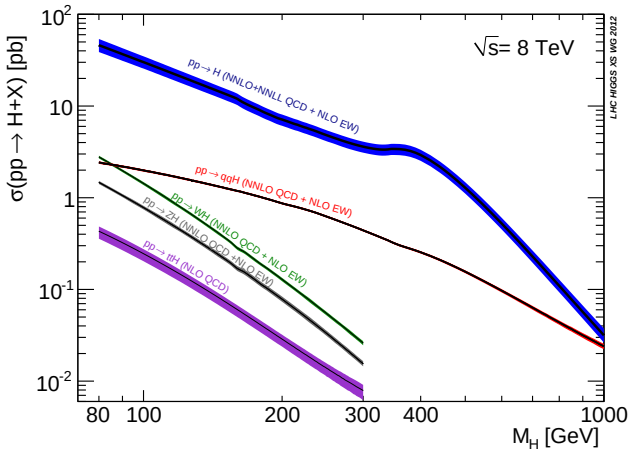

Figure 1.2: Calculated Higgs production cross sections for centre of mass energy, $\sqrt{s} = 8$ TeV [20]. ggF is denoted here by pp→H, VBF by pp→qqH and VH by pp→WH/ZH.

observed.

While ggF has the largest predicted cross section, this production process provides no additional features.This means it can be difficult to identify Higgs events against the background of Standard Model processes with similar signatures and significantly larger cross sections.

1.9.2 Vector Boson Fusion

VBF is the largest quark initiated Higgs production process. The two incoming quarks each radiate a $W^{\pm}$ or Z^0 boson and these fuse to produce the Higgs. The quarks themselves form high-p_T jets in opposite halves of the detector. These jets can be exploited to help identify events containing a Higgs produced via VBF.

1.9.3 Vector Boson Associated Production

In VH, the third and final production mechanism considered here, a $W^{\pm}$ or Z^0 is produced from two quarks. This vector boson then radiates a Higgs before decaying itself. The decay products of the $W^{\pm}$ or Z^0 may be used to help detect VH produced Higgs events. This is particularly useful in cases where the vector boson decays leptonically.

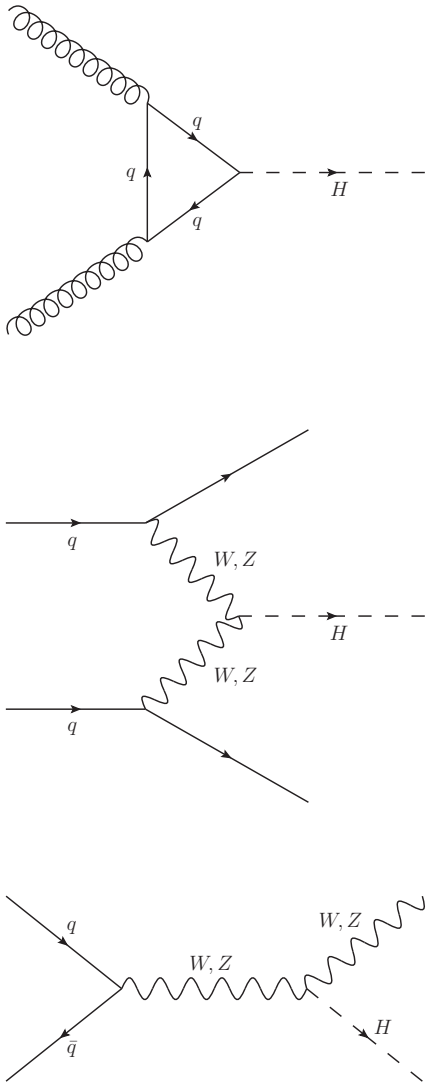

Figure 1.3: Feynman diagrams showing the three Higgs production processes considered in this analysis [21]. From top to bottom these are gluon-gluon fusion (ggF), vector boson fusion (VBF) and vector boson associated production (VH).

Chapter 2

The Large Hadron Collider

2.1 Introduction

The Large Hadron Collider (LHC) is the world's largest particle collider, with a circumference of 27 km. It is located beneath the Swiss-French border at the European Organization for Nuclear Research (CERN), near the city of Geneva. The LHC is a proton-proton collider with a design energy of 7 TeV per proton beam. This gives a total design centre of mass energy ($\sqrt{s}$) of 14 TeV.

The protons accelerated within the machine begin as hydrogen atoms. The hydrogen atoms are stripped of their electrons using an electric field, isolating protons. These protons then enter the LHC's acceleration chain. The LINAC2 linear accelerator is the first component, accelerating the protons using radio frequency (RF) cavities [22]. LINAC2 has been in operation since 1978 and accelerates the protons to an energy of 50 MeV. With the protons only possessing a minute fraction of the desired collision energy, a number of additional acceleration steps are required before injection into the LHC ring.

Next comes the Proton Synchrotron Booster (PSB) [23]. This is the first of four synchrotrons in the acceleration chain. Synchrotrons are circular accelerators which use RF cavities for acceleration as well as dipole magnets to guide the beam on a circular trajectory. Unlike linear accelerators, the fact that the beam is travelling in a circle means that each RF cavity in a synchrotron may repeatedly accelerate the protons, achieving large increases in energy. With the acceleration gradient of current RF cavities, linear accelerators are unable to accelerate protons to the energies required at the forefront of particle physics. Synchrotron acceleration, however, is limited by the radiative losses of charged particles undergoing radial acceleration (bending) due to the emission of synchrotron radiation. The PSB accelerates the

50 MeV protons from LINAC2 up to energies of 1.4 GeV.

The protons are then passed to a larger synchrotron ring named the Proton Synchrotron (PS) [24]. The PS began operation in 1959, at which point it was the highest energy particle accelerator in the world. With a circumference of 628 m, the PS is able to further accelerate the beam to an energy of 25 GeV using its RF cavities.

The final synchrotron used for acceleration before injection into the LHC is the Super Proton Synchrotron (SPS) [25]. In 1983, colliding protons and antiprotons, this machine was responsible for the discovery of the W and Z bosons. The SPS accelerates the protons of the beam to 450 GeV, the injection energy of the LHC.

The SPS is used to alternately fill the LHC in a clockwise and anticlockwise direction. The energy of the two beams is then ramped up to collision energy using the LHC's 16 RF cavities [26]. Along the circumference of the LHC there are over 1,000 dipole magnets, used to bend the beam into a circle. There are also approximately 400 quadrupole magnets, which are used to focus the protons into a beam. Due to the extremely high currents required to maintain the magnetic fields, the magnets of the LHC are superconducting electromagnets. This means they must be kept at a temperature below 2 K using liquid helium. To minimise interference with the proton beam, the two beam pipes must be kept in a state of vacuum.

Once in the LHC, the protons of the beam are grouped in discrete bunches, typically around 1,000 or so. The two beams are able to cross at four different locations around the ring, each crossing point home to a cutting edge particle detector. The bunches of the two beams cross 40 million times per second at each crossing point.

The two beams of protons are so tightly focused that each bunch crossing results in multiple proton-proton interactions. The additional interactions are a background activity to any physically interesting processes. This phenomenon is referred to as pile-up and the mean number of interactions per bunch crossing for the ATLAS detector can be seen in Figure 2.1. The high frequency of bunch crossings and multiple interactions per crossing are vital as many of the processes the LHC's experiments are searching for occur extremely rarely.

The LHCb (Large Hadron Collider beauty) [28] experiment is designed to focus on particle collisions where b-quarks are produced. These are especially interesting as they may shine a light on the matter-antimatter asymmetry observed in the Universe. Alongside more typical detector components, LHCb is equipped with two specialist components which help it accurately identify events containing

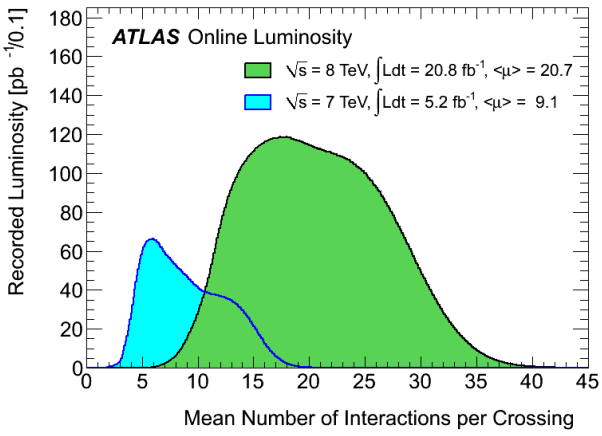

Figure 2.1: Luminosity-weighted distribution of the mean number of interactions per crossing for the 2011 and 2012 data [27].

b-quarks. These are a very precise vertex locater (used to identify the exact position of the collision) and a Cherenkov detector (which helps distinguish between different hadrons).

ATLAS (A Toroidal LHC ApparatuS) [29] and CMS (Compact Muon Solenoid) [30] are two multi-purpose detectors located on opposite sides of the LHC ring. They are capable of detecting all of the particles of the Standard Model, except for the rarely interacting neutrinos, and have detector coverage approaching 4π steradians. The two experiments use this ability to search for never before seen physics as well as making precision measurements of known interactions.

The fourth collision point is home to ALICE (A Large Ion Collider Experiment) [31]. As well as being filled with protons, the LHC can also be used to accelerate lead ions. ALICE is used to study the high energy densities produced in lead-lead collisions. These studies include investigations into quark-gluon plasmas, a phase of matter similar to that of just after the Big Bang.

2.2 Operational History

Construction of the LHC and its associated particle detectors was completed in June 2008. By August of that year, the beampipe had been cooled to the operating temperature of 1.9 K and was in vacuum. The first beams of protons were circulated individually through the two LHC beampipes in early September.

Fewer than two weeks later, on the 19th September 2008, one of the superconducting magnets of the LHC quenched [32]. A quench occurs when the magnetic

field or temperature of the magnet becomes high enough for the magnet material to leave the superconducting state. Rapid heating occurred causing around two tonnes of the liquid helium used to cool the magnets to vent explosively. The large explosion damaged over 50 of the magnets and coated parts of the beampipe in soot. It was ultimately caused by a high current arcing between two magnets due to imperfect soldering between the two.

Over the following year, repairs were made to the accelerator. The damaged magnets were repaired or replaced and other faulty connections were identified and repaired. Following the incident, beams were again circulated through the machine in November 2009. The first proton-proton collisions at a centre of mass energy of 450 GeV occurred later that month.

With the machine operating smoothly, there was just one week between the first collisions and the LHC claiming the world record for collisions at the highest centre of mass energy. On the 30th November 2009, the record was set at $\sqrt{s} = 2.36$ TeV, beating the record set years earlier by Fermilab's Tevatron.

In the following months, the record was repeatedly broken before reaching the planned centre of mass energy of 7 TeV in March 2010. This value was lower than the design energy as some solder connections in the LHC still required replacing. The upgrade was scheduled for the long shutdown of the machine following around three years of data taking.

After the success of data taking at $\sqrt{s} = 7$ TeV in 2011, for 2012 the centre of mass energy was increased to 8 TeV. This change slightly increased the chance of a magnet quenching but the increased risk was deemed acceptable when compared to the benefits. One of the principal benefits was the increase in cross section for many physically interesting processes, including Higgs production.

During these two years of operation, the ATLAS experiment recorded integrated luminosities of 4.5 fb^{-1} at $\sqrt{s} = 7$ TeV and 20.3 fb^{-1} at $\sqrt{s} = 8$ TeV. This forms the dataset used in this analysis.

Following the data taking of 2012, the LHC and each of the experiments were shutdown for a period of approximately two years. This shutdown was used to fix or improve various components of the accelerator and detectors. One key repair is the resoldering of the final potentially faulty superconducting magnet connections. This should allow the LHC to run at energies closer to its design energy in the future.

Chapter 3

The ATLAS Detector

3.1 Introduction

The ATLAS detector is one of two multipurpose detectors at the LHC, the other being CMS. Multipurpose experiments are required to have good directional coverage as well as sensitivity to as wide a range of particles as possible. To achieve these goals, the ATLAS detector is made up of multiple subsystems which each target different classes of particles. These subsystems, displayed in Figure 3.1, will be introduced now but discussed in more detail below.

The magnet system generates a magnetic field which envelops the entirety of the detector. As charged particles follow curved trajectories in a magnetic field, information on the charge and momentum of any charged particles in the detector can be inferred. It consists of a solenoidal component which produces a magnetic flux in the inner detector, and a toroidal component which induces a flux in the muon system.

At the very centre of the ATLAS detector lies the LHC beampipe. The protons of the two beams travel in opposite directions in two pipes, crossing at the centre of ATLAS so proton-proton collisions can occur. The beampipe's shielding is reduced within the ATLAS detector to allow detector components to be located closer to the interaction point (IP).

The first layer of the detector is the inner detector (ID). The inner detector is capable of determining the tracks of charged particles to a very high precision. This is crucial for both determining which particles in each event originate from the collision vertex as well as whether any particles were produced in secondary decays. Measuring the curvature of tracks in the inner detector's multiple layers also provides information on the momenta and charges of any particles produced.

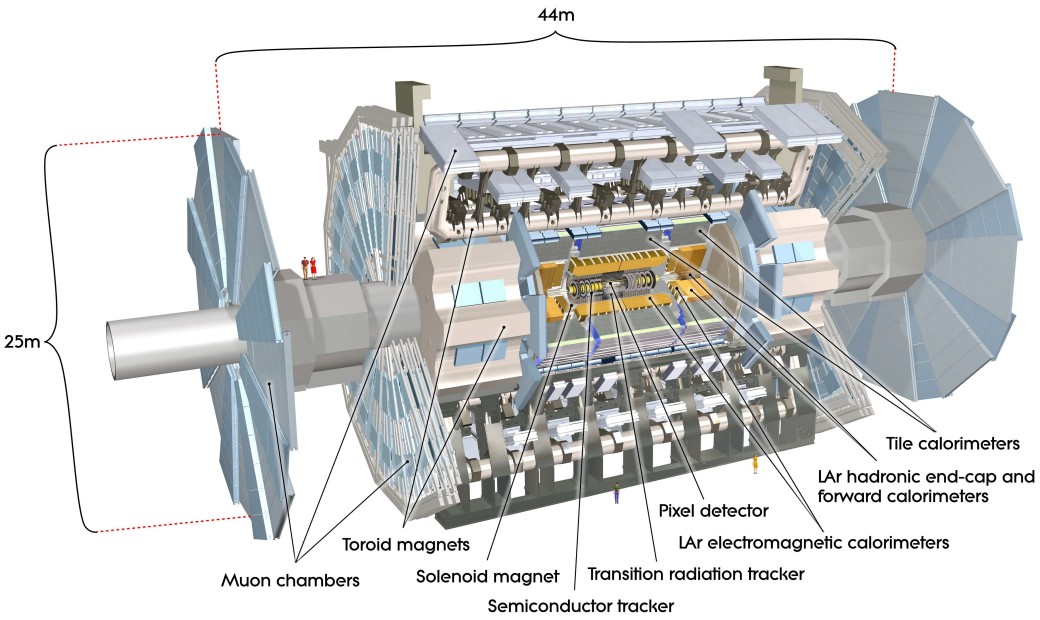

Figure 3.1: Cutaway view of the ATLAS detector, with the major subcomponents labelled [29].

The electromagnetic calorimeter (ECAL) surrounds the inner detector. The purpose of a calorimeter is to absorb incident particles, measuring their energy in the process. The ECAL is designed to be most sensitive to electrons and photons.

This is followed by the hadronic calorimeter (HCAL). The HCAL is used to measure the energy of hadrons such as protons, neutrons and pions. It can also be used to measure jet energies.

The final major subdetector is the muon system. Muons are capable of penetrating both calorimeters so dedicated subdetectors are installed surrounding the HCAL. These detect the passage of muons, allowing their properties to be measured (in conjunction with the inner detector and the muon system's toroidal magnets).

3.2 The ATLAS Coordinate Systems

Before proceeding with more detailed specifications of each subdetector, the coordinate systems used by the ATLAS collaboration must be described. The ATLAS detector consists of a barrel with an endcap on either end, making it cylindrical in shape. Both a Cartesian and cylindrical coordinate system is in use, each useful for different applications. The origin of each coordinate system is defined as the nominal interaction point, at the centre of the ATLAS detector.

The Cartesian system defines the positive x direction as being towards the centre of the LHC ring. The positive y direction is defined as being straight up. In keeping with a right-handed coordinate system, this leaves the z-axis to be defined in the anti-clockwise direction if viewing the LHC ring from above.

The z-axis is used to define two halves of the detector. The half of the detector where z is positive is named the A-side with the other half being called C-side.

The cylindrical coordinate system uses the same z-axis definition as the Cartesian one. ϕ is defined as the angle around the beampipe, with $\phi = 0$ corresponding to straight up (along the y-axis). ϕ values run from 0 to 2π. Pseudorapidity, η, is used as the third and final coordinate. With θ defined as the angle from the positive z-axis, η is defined by

$$\eta = -\ln(\tan(\theta/2)). \tag{3.1}$$

This definition means any particle located on the transverse plane through the origin will have an η of 0. The η of locations away from this plane increase as they approach the beampipe, tending towards infinity.

Two variables which may be used to parameterise a tracks closest approach to the nominal interaction point are the impact parameters of the track. The transverse impact parameter, d_0, is the distance of closest approach between the track and the z-axis. The z-coordinate of this point is named the longitudinal impact parameter, z_0.

While discussing coordinate systems, one final important definition is required. The angular distance, ΔR, between two coordinates is defined as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{3.2}$$

$\Delta\eta$ and $\Delta\phi$ are the angular separations in the η and ϕ planes respectively.

3.3 Magnet System

The magnet system of the ATLAS experiment consists of four distinct elements. These are the barrel solenoid, the barrel toroid and the two endcap toroids. Each is a superconducting electromagnet and, when operating at their design currents, represent a combined total stored energy of 1.6 GJ [29]. The conducting wires are manufactured from a niobium/titanium/copper mix, stabilised with aluminium.

To minimise the uncertainty on momentum measurements of any charged

particles, great care is taken to measure the magnetic field to a very high precision.

3.3.1 Barrel Solenoid

The barrel solenoid lays between the inner detector and calorimeters, providing a near uniform 2 T field for the inner detector. Its location allows it to use the steel of the hadronic calorimeter as a return yoke for its magnetic flux. It forms a cylinder centred on the beampipe, with an inner diameter of 2.46 m and an outer one of 2.56 m. The cylinder is 5.8 m in length.

Due to its location inside the calorimeters, care was taken when designing the solenoid to minimise the material thickness. The final thickness was measured to be roughly 0.66 radiation lengths at normal incidence.

A current of 7.73 kA is required to reach the design magnetic field strength, giving a total stored energy of 40 MJ [29]. The nominal temperature for the superconducting magnet is 4.5 K but the magnet can safely quench by heating up the surrounding detector components. To reach and maintain these temperatures, the solenoid shares a cryostat with the electromagnetic calorimeter which surrounds it. This sharing also helps achieve the design goal of minimising the material thickness within the ECAL, by eliminating the need for two vacuum walls.

To ensure the magnetic field strength is not the dominant uncertainty on momentum measurements using the inner detector, the fractional bending power must be measured to a precision of around 1 part in 2,000. This target was achieved by mapping the magnetic field using dedicated equipment once the inner detector had been fully constructed [33].

3.3.2 Toroids

Both the barrel and endcap toroids consist of eight identical loops, distributed symmetrically. The toroids are responsible for inducing a magnetic field in the muon detectors, allowing muon momenta to be calculated. This measurement can be combined with the corresponding one from the inner detector, providing an improved muon momentum resolution.

The barrel loops are over 25 m in length, with inner and outer diameters of 9.4 and 20.1 m respectively. These loops give a 0.5 T magnetic field in the muon detector, requiring a current of 20.5 kA [29]. Again, the stored energy of 1.1 GJ can safely be quenched into the surrounding material without a damaging rise in temperature. The total weight of the barrel toroid is around 830 tonnes.

The two endcap toroids are identical in design. They each feature eight

loops held in place by steel keystones between them. The endcap toroids have inner diameters of 1.65 m and external diameters of 10.7 m. They are 5 m long.

The same 20.5 kA current as used in the barrel toroid, induces a magnetic field of 1 T in the endcap toroids. Each endcap weighs around 240 tonnes in total.

To achieve a momentum measurement uncertainty below 5 %, the magnetic fields in the muon detectors must be measured to an accuracy of at least 1 in 200. Due to the large volume of the toroids, manual measurement of the field would be impractical. 1,840 B-field sensors are installed in the toroids to allow, in conjunction with simulations, the field to be determined sufficiently accurately [34].

3.4 Inner Detector

The inner detector (ID) surrounds the interaction point and is responsible for precision momentum measurements and vertex locating. Its position makes it sensitive to both primary vertices (where the proton-proton interaction occurred) and secondary vertices (from heavy-flavour or tau decays).

The location of the ID means it is exposed to a very high level of radiation throughout its ten year life. This also means the multiplicity of tracks in the ID will be extremely high, therefore the sensors must be of very high granularity.

In order to meet these design goals while still keeping costs manageable, the ID is made up of three independent subsystems. The first layers form the silicon pixel tracker, a detector made of extremely high granularity semiconductor pixels and able to cope with the high track multiplicity.

Surrounding this, where the track multiplicity is slightly reduced, is the silicon microstrip tracker (SCT). This is similar in design to the pixel tracker, albeit with a significantly reduced granularity to keep costs acceptable. The two silicon trackers provide tracking information in the range $|\eta| < 2.5$.

The third and final subsystem is named the transition radiation tracker (TRT). It uses approximately 350,000 drift tubes to measure the paths of charged particles. While roughly ten times less precise than the silicon trackers, it provides many more tracking hits so significantly aids momentum measurement precision. Identification of electrons is also made possible by detecting transition radiation produced by incident particles. The TRT tracks particles with $|\eta| < 2.0$.

For the combined ID, the impact parameter resolution is of the order 0.02 mm for d_0 and 0.1 mm for z_0 [35]. The intrinsic efficiencies are almost 100% for the silicon detectors and around 97% for the TRT.

An overview of the ID can be seen in Figure 3.2. A more detailed view is

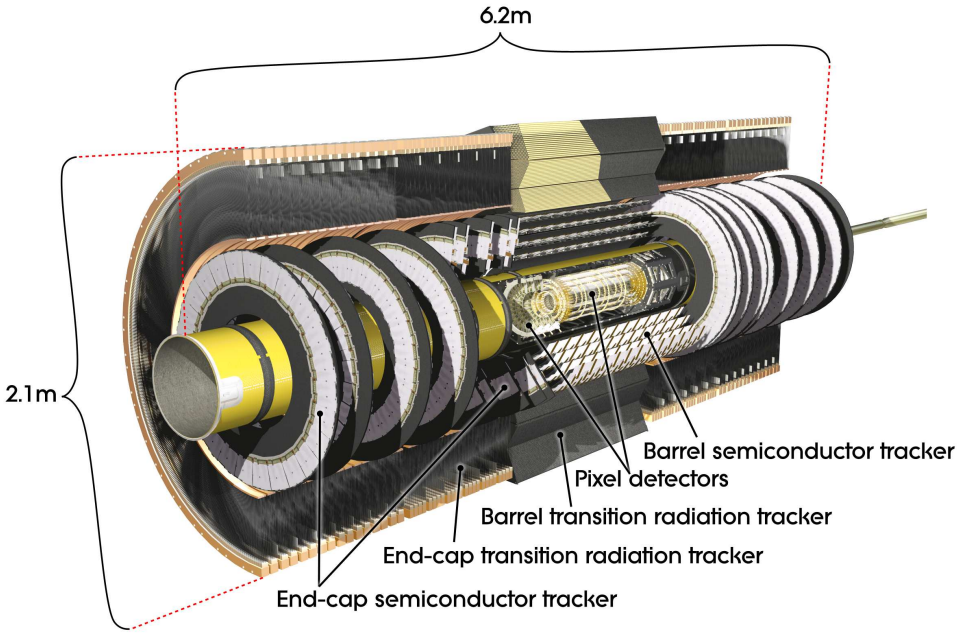

Figure 3.2: Cutaway view of the ATLAS inner detector, showing the locations of the pixel tracker, SCT and TRT [29].

shown in Figure 3.3.

3.4.1 Silicon Pixel Tracker

The pixel tracker has three layers in the barrel and each endcap. The barrel layers are at radii of 50.5, 88.5 and 122.5 mm and stretch to $|z| = 400.5$ mm. The endcaps are located at z positions of 495, 580 and 650 mm and form circles with inner and outer radii of 88.8 and 149.6 mm respectively. This layout ensures any particle with $|\eta| < 2.5$ will cross at least three layers of the pixel tracker.

A bias voltage of 150 V is applied to the silicon semiconductor in each pixel. This bias can be safely increased to 600 V, to maintain performance after the expected radiation damage occurs.

In total, the pixel tracker is made up of 1,744 sensor modules. A typical sensor contains nearly 50,000 pixels and has dimensions of 50×400 μm. In the barrel the pixels are orientated with the long edge parallel to the beampipe whereas the endcap pixels are placed radially. Each pixel is 250 μm thick. These properties mean the pixel detector has an $R - \phi \times z$ ($R - \phi \times R$ in the endcaps) hit resolution of 10×115 μm [29].

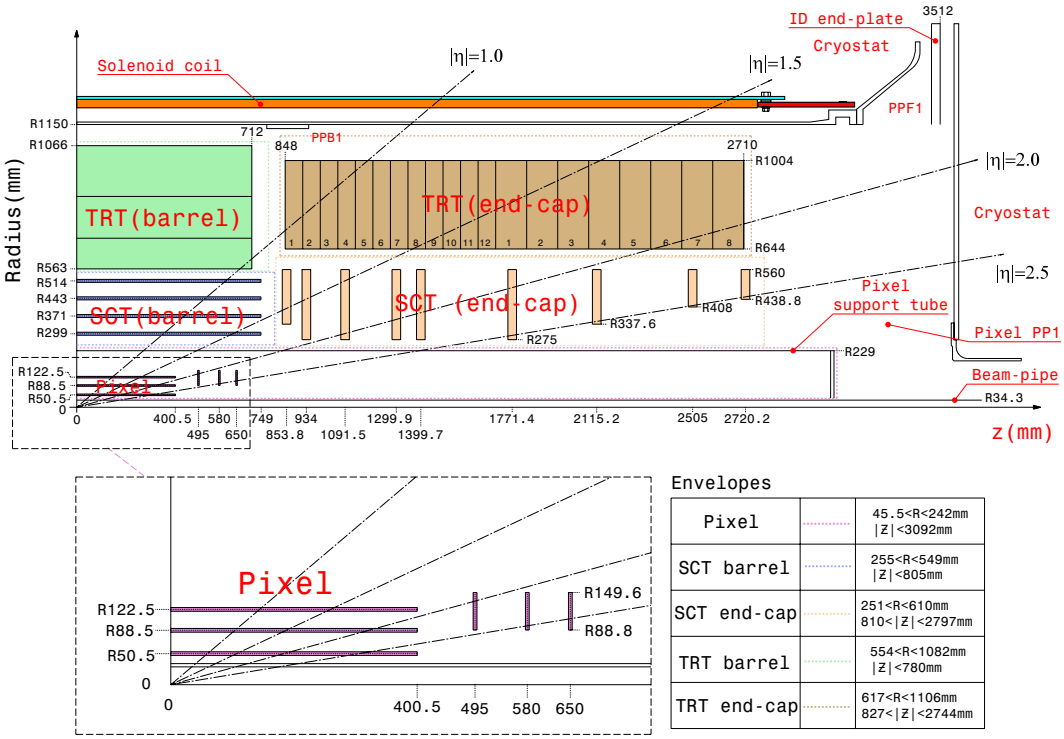

Figure 3.3: Cross section through a quarter of the ATLAS inner detector [29]. The locations and dimensions of each subdetector can be seen.

3.4.2 Silicon Microstrip Tracker

The SCT is similar in design to the pixel tracker, but with a reduced granularity and hence resolution. A typical silicon strip of the SCT has dimensions of 80 μm $\times$ 12 cm, orientated in the same way as the silicon pixels. Each strip is slightly thicker than a pixel, at 285 μm. In total there are nearly 16,000 sensors, each with around 800 strip modules.

The SCT is arranged into four barrel layers and nine disks in each endcap. Each layer is double sided and, to improve the resolution along the z-axis, one side is offset by a stereo angle of 40 milliradians. The barrel layers are at radii of 299, 371, 443 and 514 mm and continue to $z = 749$ mm. The nine endcap layers are spread between $z = 853.8$ mm and $z = 2{,}720.2$ mm The layers are positioned so that any particle with an η under 2.5 will hit layers of the SCT.

These properties give the SCT a final hit resolution of 17×580 μm [29].

3.4.3 Transition Radiation Tracker

The TRT is a straw tracker, made up of over 350,000 straws of diameter 4 mm. Straws parallel to the beampipe form the barrel TRT, occupying the volume $563 < R < 1,066$ mm, $|z| < 712$ mm. In the endcap, the straws radiate out from the beampipe with z values ranging from 848 to 2,710 mm. The nature of a straw tracker means the TRT only provides a position measurement in one dimension.

The wire in each straw is made from a 31 μm diameter strand of pure tungsten, plated with a thin layer (< 0.7 μm) of gold. The straws themselves are primarily made of 25 μm thick polyimide, covered with a 0.2 μm coating of aluminium and a roughly 5 μm layer of graphite-polyimide. The anode wires are held at ground voltage and the voltage of each straw is $-1,530$ V during normal operation.

The nominal gas mix in the TRT is 70% xenon, 28% carbon dioxide and 3% oxygen. Transition radiation, produced when a charged particle crosses from one material to another, ionises the TRT gas and hence produces a signal in the straw. Due to the inverse mass dependence of transition radiation, electrons produce a much larger signal than the more massive minimum-ionising charged particles. This fact means the TRT forms an important component of the ATLAS detector's electron identification capabilities [29].

Each straw in the barrel runs the full 1.44 m length of the barrel TRT. The wire in each barrel straw has a 6 mm break in the centre of the detector (at $z = 0$ mm). While necessary to reduce the occupancy of the straws, this leads to an inefficiency for particles crossing with $|z| < 20$ mm. The straws of the endcap are

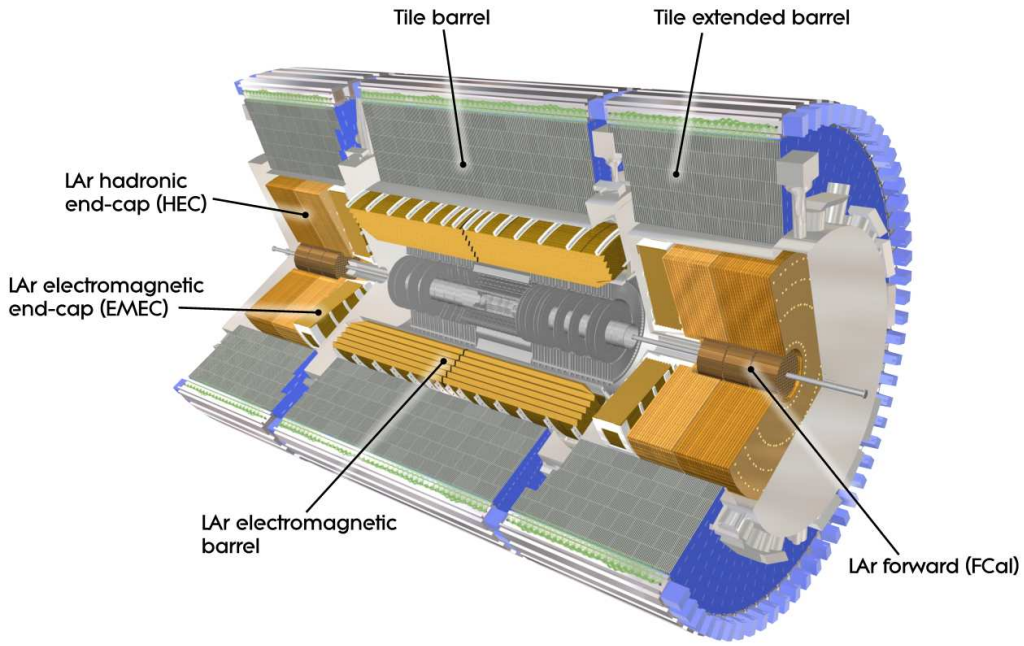

Figure 3.4: Cutaway view of the ATLAS calorimeters, showing the major subcomponents of each [29].

0.37 m in length, radiating outwards from $R = 644$ mm.

Limited by the drift time accuracy, the resolution of the TRT straws is approximately 130 μm [29]. Despite the comparatively lower resolution and the fact the TRT only provides a measurement in the $R - \phi$ coordinate, the large number of hits mean the TRT meaningfully contributes to momentum measurements. A particle with $|\eta| < 2.0$ will typically leave 36 hits in the TRT.

3.5 Calorimetry

Calorimeters are a class of particle detector which measure the energy of incident particles by absorbing them. The ATLAS detector has two calorimeters, the electromagnetic calorimeter and the hadronic calorimeter, and the layouts of both can be seen in Figure 3.4. Both of the calorimeters are sampling calorimeters.

A sampling calorimeter is made up of alternating layers of absorber and active material. The absorber is a dense material which absorbs some of the energy of any particles passing through. When traversing the active material layers, the incident particles cause secondary particles to be emitted. Detecting these secondaries allows the energy of the original particle to be measured.

Good containment of incoming particles is a key feature of a good calorimeter. This allows the energies of all particles, and the missing transverse energy (E_T^{miss}), to be measured accurately. The E_T^{miss} is a measure of the energy imbalance in the calorimeters, due to particles escaping detection, and is explained in detail in Section 5.2.6) The ATLAS calorimeters have full coverage in ϕ and η coverage up to $|\eta| = 4.9$. The two calorimeters are also sufficiently thick. The electromagnetic calorimeter is greater than 22 radiation lengths (X_0) deep and the hadronic calorimeter's depth is more than 10 interaction lengths (λ) throughout.

Combined, these properties mean the calorimeters can measure the energies of electrons, photons, hadrons and jets, as well as the E_T^{miss}, to a high accuracy and precision. The properties of the individual calorimeters are discussed in detail below.

3.5.1 Electromagnetic Calorimeter

The ATLAS electromagnetic calorimeter is a sampling calorimeter which uses lead as its absorber and liquid argon (LAr) as the active material. The density of the lead makes it a good absorber material and the ionisation properties and radiation hardness of the LAr make it an appropriate choice for the active layers. The absorber plates range in thickness, depending on η, from 1.1 to 2.2 mm and are strengthened with a thin (0.2 mm) layer of steel on each side.

The calorimeter has an accordion shaped geometry meaning particles typically cross the active layer at an angle. This has the advantage of increasing the path length through the active material while not increasing the drift distance for the produced secondary electrons.

As well as the usual barrel and endcap split, the EM calorimeter has a number of additional components. The barrel covers the range $|\eta| < 1.475$ and the endcap $1.375 < |\eta| < 3.2$. To improve coverage there is an additional forward calorimeter (FCal) covering the range $3.1 < |\eta| < 4.9$.

The final significant component is the presampler. This helps measure the energy lost before particles reach the EM calorimeter for particles with $|\eta| < 1.9$. The presampler is 11 mm and 5 mm thick in the barrel and endcap region respectively. It is made up entirely of LAr, the active material of the EM calorimeter, and measures showers produced before the EM calorimeter.

One further distinction is made, namely the precision measurement region (PMR). This refers to the area of the EM calorimeter which overlaps with the inner detector, where an improved accuracy and precision is required compared to the rest of the calorimeter.

For a 20 GeV electron, the fractional energy resolution of the EM calorimeter is approximately 2% [29]. This decreases as electron energy increases, reaching 1% for a 100 GeV electron.

Barrel Electromagnetic Calorimeter

The barrel is split into two halves, with a 4 mm gap between them at $z = 0$ mm. Each half is 3.2 m long with an inner and outer radius of 1.4 and 2 m respectively. This thickness means the minimum number of radiation lengths traversed by an incident particle is 22 in the barrel. Coverage by the barrel is provided up to $|\eta| = 1.475$ and the weight of each half is 57 tonnes.

The PMR, which encompasses the entire barrel as well as some of the endcaps, has three separate layers of calorimetry. The second layer, which forms the bulk, is subdivided into radial towers with $\Delta\eta \times \Delta\phi$ measurements of 0.025×0.025 [29]. The third layer in the barrel has tower dimensions of 0.050×0.025.

Towers in the first layer are less finely segmented in ϕ than the second but are eight times more finely segmented in η. This additional segmentation improves the EM calorimeter's η resolution when measuring photons, as well as its ability to distinguish between photons and neutral pions ($\pi^0 \to \gamma\gamma$).

At high barrel $|\eta|$s (greater than 1.35) there are only two, less finely segmented, layers in the barrel. This is due to geometric constraints and overlap with the endcap acceptance. In total, the barrel EM calorimeter has over 100,000 readout channels.

Endcap Electromagnetic Calorimeter

Each of the endcap EM calorimeters are 0.63 m thick and weigh 27 tonnes. They occupy the volume with coordinates $1.375 < |\eta| < 3.2$, $330 < R < 2,098$ mm. Each endcap is split into two concentric rings at $|\eta| = 2.5$.

The outer ring ($1.375 < |\eta| < 2.5$) varies in segmentation as a function of η. Between $|\eta| = 1.5$ and $|\eta| = 1.8$ the endcap calorimeter has three layers, segmented identically to the bulk of the barrel calorimeter [29].

Similar to the barrel, in the overlap region ($1.375 < |\eta| < 1.5$, where the barrel and endcap calorimeters overlap) there are only two, more coarsely segmented layers. The granularity also decreases for higher η values, where precise energy measurements are less vital. The inner ring of the endcap ($2.5 < |\eta| < 3.2$) has two uniform layers with tower dimensions of 0.1×0.1 ($\Delta\eta \times \Delta\phi$).

Each endcap has a total of over 30,000 readout channels.

3.5.2 Hadronic Calorimeter

The hadronic calorimeter is split into three major components. These are the barrel hadronic calorimeter, or tile calorimeter, the hadronic endcap (HEC) and the forward calorimeter (FCal). The three will be discussed in more detail individually below.

Each of the three subsystems are sampling calorimeters, although they use different absorbers and active materials. The hadronic calorimeter provides $|\eta| < 4.9$ and near full ϕ coverage.

Tile Calorimeter

The tile, or barrel calorimeter, is itself split into three separate pieces. The 5.8 m long central barrel and the two 2.6 m long extended barrels on either end. Each of the barrel components have inner and outer radii of 2.28 and 4.25 m respectively. These dimensions mean the central piece covers $|\eta| < 1.0$ while the extended barrel has a slightly overlapping η coverage of $0.8 < |\eta| < 1.7$.

The tile calorimeter uses steel as its absorber, while the active material is made from scintillating tiles. The scintillating tiles are placed radially which allows near seamless coverage. Wavelength shifting fibres connect the scintillator to photomultiplier tubes (PMTs) for readout of the detector. The ratio of absorber to active material is around 5:1.

Each of the three barrel components are split into 64 individual modules by ϕ. These modules each have three layers of varying thicknesses. In the central barrel (extended barrel) the three layers have thicknesses of 1.5, 4.1 and 1.8 (1.5, 2.6 and 3.3) nuclear interaction lengths (λ).

The first two layers have a $\Delta\eta \times \Delta\phi$ granularity of 0.1×0.1 [29]. The third layer is twice as coarse in η.

The tile calorimeter includes additional instrumentation in the gap between the central and extended barrel components. This is designed to partially recover energy lost in the region.

The fractional energy resolution for the tile calorimeter is around 14% for a 20 GeV incident particle [29]. This drops to approximately 6% for a 350 GeV particle.

Hadronic Endcap

The absorber and active material of the HEC are copper and LAr respectively. As liquid argon needs to be kept cool, the HEC shares a cryostat with the EM

calorimeter endcap. The HEC covers an η range of $1.5 < |\eta| < 3.2$, meaning it overlaps with both the tile calorimeter and FCal.

Each of the endcaps is made up of two "wheels", giving four wheels in total. The wheels are divided in ϕ into 32 wedge shaped modules. The outer radius of each of the four wheels is 2.03 m.

The first wheel on each side, HEC1, is made up of 24 25 mm thick absorber plates. LAr occupies the 8.5 mm gaps between each of the 24 layers. These parameters give HEC1 a sampling fraction of 4.4%.

There are 16 layers of absorber in the secondary wheels (HEC2), with each absorbing layer being twice as thick as the layers in HEC1. HEC2 has a 2.2% sampling fraction. The HEC has a $\Delta\eta \times \Delta\phi$ granularity of 0.1×0.1 for $|\eta| < 2.5$ and 0.2×0.2 outside that region [29].

Forward Calorimeter

The FCal, like the HEC and EM calorimeter, uses LAr as its active material. It covers the η range $3.1 < |\eta| < 4.9$. The absorber used depends on the layer of the FCal.

The first of the three 0.45 m layers uses copper as its absorber. This layer is optimised for sampling electromagnetic showers. The second and third layers' absorbers are made from tungsten. These are optimised for hadronic showers, with the tungsten providing good shower containment.

Being so close to the beampipe, the FCal experiences very high particle fluxes. This means that the active material must only make up a small fraction of the calorimeter.

The LAr is placed in cylinders which run parallel to the beampipe. The cylinder walls, where the LAr is located, are around 0.2–0.5 mm thick depending on the FCal layer. The volumes of the cylinders are filled with the absorber material.

In total, the FCal provides around 10 interaction lengths (λ) of absorption. A copper shielding plug is placed behind the third layer of the FCal to further reduce punch through from the calorimeter to the muon systems behind.

3.6 Muon System

The primary performance goal of the muon system is the ability to measure the p_T of a 1 TeV muon with a better than 10% relative uncertainty [29]. Such a muon has a sagitta of approximately 500 μm, meaning the muon system needs a resolution better than 50 μm in the bending plane of the magnets. This requirement means

the muon system is able to provide a standalone measurement of muons in the p_T range 3–3,000 GeV.

The variety of flux conditions which the muon system must cope with, as well as the precision measurement and triggering needs, mean the muon system requires four different subcomponents.

Two of the four are designed to precisely measure track bending, meaning a good p_T resolution can be achieved. Monitored drift tubes (MDTs) achieve this goal throughout the whole η range covered by the muon system. These are complemented by the cathode strip chambers (CSCs) which are better able to cope with the high particle flux at high-η ($|\eta| > 2.0$) in the first endcap wheel.

The two remaining subdetectors must have a fast response (tens of nanoseconds), to allow the ATLAS detector to trigger on muon events. In the barrel, monitored drift tubes (MDTs) are used for event triggering. The endcap trigger requires an improved granularity and radiation hardness. The use of thin gap chambers (TGCs) satisfies these performance requirements.

Integrated with the muon system are the three toroidal magnets of the ATLAS experiment. Measuring how the tracks of muons bend in the magnetic field is key to determining the momenta of the muons. Similar to the toroidal magnets, the muon system follows an eight fold symmetry in ϕ. However, the presence of the supporting feet of the magnet system mean the muon acceptance is significantly reduced in two barrel regions.

The barrel toroid covers the range $|\eta| < 1.4$ and provides a bending power which varies between 1.5 and 5.5 mT. The two endcap toroids give a bending power of 1–7.5 mT over the η range $1.6 < |\eta| < 2.7$. A lower bending power is present in the transition region ($1.4 < |\eta| < 1.6$) where the fields of the barrel and endcap toroids overlap.

To achieve the required measurement precision, the various parts of the muon system must be carefully aligned. An optical alignment system was installed to help achieve this. Furthermore, the magnetic field must be modelled to better than a few parts in a thousand. 1,800 Hall sensors monitoring the field mean this requirement can be met.

Each of the four subcomponents is briefly described below. The layout of the four, alongside the toroids, can be seen in Figures 3.5 and 3.6.

3.6.1 Monitored Drift Tubes

The MDTs form the bulk of the muon system's precision track measuring ability, covering the full range of $|\eta| < 2.7$. They form three cylinders in the barrel and

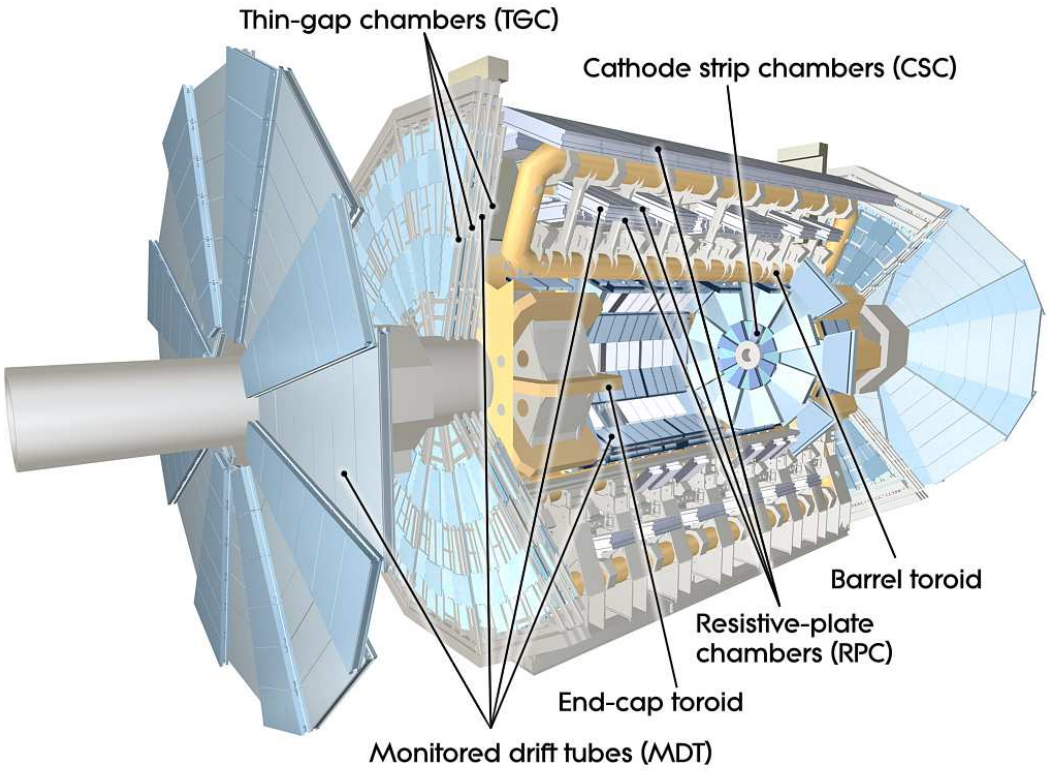

Figure 3.5: Cutaway view of the ATLAS muon system, showing the toroidal magnet systems and the four muon subdetectors [29].

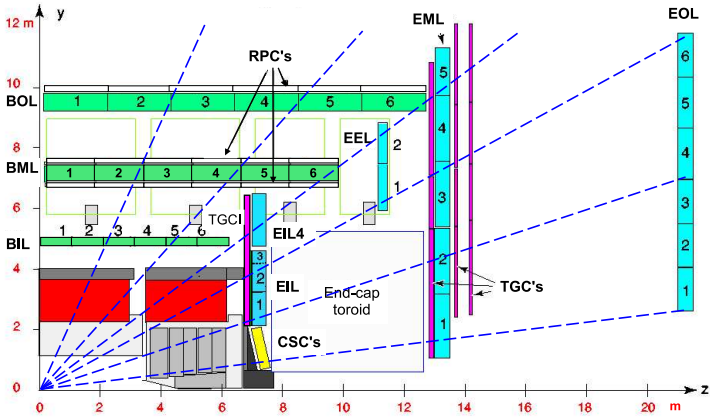

Figure 3.6: Cross section showing the locations and dimensions of the components of the ATLAS muon system [29].

three wheels in each endcap, giving a total surface area of 5,500 m^2.

The three barrels have radii of approximately 5, 7.5 and 10 m, with each extending to an $|\eta|$ of 2.5. The endcap wheels are located at $|z| = 7.4, 14, 21.5$ m. Additional MDTs are placed between the first and second endcap wheels ($|z| = 10.8$ m), to guarantee any straight track passes through three layers.

Each layer of MDTs contains 3–8 rows of drift tubes. The MDTs are simple in construction and provide a high accuracy for track measurements. With each tube giving an individual spatial resolution of 80 μm, a chamber gives a resolution of around 35 μm [29]. MDTs only provide a measurement in one coordinate, meaning the second coordinate must be supplemented by the RPCs/TGCs.

3.6.2 Cathode Strip Chambers

Due to the higher particle flux, CSCs replace the MDTs in the first wheel of the endcap for $2.0 < |\eta| < 2.7$. As a form of multiwire proportional chamber (MWPC), their improved time resolution allows them to cope with the increased flux. The cathodes are split into strips orthogonal to the anode wires, allowing the first and second coordinates to be measured.

The CSCs have a resolution of 40 μm in the R (bending) coordinate and 5 mm in the ϕ coordinate [29]. The total CSC surface area is approximately 65 m^2.

3.6.3 Resistive Plate Chambers

The RPCs are the first of the two muon system subcomponents whose primary purpose is triggering. They are located in the barrel ($|\eta| < 1.05$), where the magnetic field is smoother and particle flux is lower than in the endcaps. These conditions mean RPCs can provide measurements of a sufficient spatial and temporal resolution.

The RPCs form three cylinders in the barrel, each cylinder providing two $\eta \times \phi$ measurements. The first two cylinders are located on either side of the second MDT layer, and the third surrounds the third MDT layer. This layout gives a final $z \times \phi$ resolution of 10×10 mm [29].

3.6.4 Thin Gap Chambers

The TGCs perform triggering duties in the endcap, covering a range of $1.05 < |\eta| < 2.4$. The instrumentation actually expands to an $|\eta|$ of 2.7 but the higher η regions cannot be used for triggering. TGCs are a form of MWPC with good time resolution and high rate capabilities.

Triggering in the endcap is challenging for a number of reasons. Firstly, closer to the beampipe, particle fluxes are higher. This means the detectors must be radiation hard, as well as having a rapid response.

Secondly, a muon in the endcap has higher total momentum for the same transverse momentum, when compared to a muon at lower η. This means high-η detectors must have an improved granularity, to measure track bending to a sufficient precision.

TGCs are able to cope with these constraints meaning they are ideal for triggering in the ATLAS endcap. The three primary TGC layers are located around the middle MDT layer in the endcaps. A fourth layer is included, attached to the inner plane of the inner MDT wheel.

The three primary wheels are located close together and outside the endcap toroid's magnetic field. The limited bending of tracks and lack of z distance between the three main subdetector layers somewhat limits the resolution of the TGCs. The TGCs have an $R \times \phi$ resolution of around 4×5 mm [29].

Chapter 4

Trigger and Data Acquisition

4.1 Introduction

The trigger and data acquisition (DAQ) systems [29] are the ATLAS components which bridge the divide between collisions occurring in the ATLAS detector and being written to permanent storage. The trigger system is responsible for inspecting the detector output from each bunch crossing, discarding uninteresting events and marking the remaining ones for storage. The DAQ performs the readout of the detector subsystems, supplying the trigger system with requested information and then saving events to disk which are approved by the trigger.

In this chapter, the operation of the trigger and DAQ system will be described. A study of the scale factors and systematic uncertainties associated with the muon trigger will also be discussed.

4.2 The ATLAS Trigger and Data Acquisition System

The trigger system's design brief is to cut the 40 MHz event rate down to a manageable output stream of 200 Hz [29]. This reduction is split into three distinct steps, with each level having more time to come to a decision than the proceeding one. Level-1 (L1) must reduce the event rate from 40 MHz to around 75 kHz. Level-2 (L2) decreases the rate further to roughly 3.5 kHz. The third and final layer, the Event Filter (EF), cuts the event frequency to the output rate of the DAQ, approximately 200 Hz. Collectively, the L2 and EF are known as the High Level Trigger (HLT).

Each level of the trigger has an associated list of trigger conditions. If an event satisfies one of the trigger conditions, commonly referred to as trigger items,

then it is accepted and passed to the next level of the trigger. The trigger items are chained together across the multiple levels of the trigger. For example, a trigger object which passes a muon trigger at L1 will only be considered by muon related triggers at the higher levels.

The full list of trigger chains is named the trigger menu. Only events passing a full chain (one trigger item at L1, L2 and the EF) are saved to disk. Each trigger item may be present in multiple trigger chains.

The most common type of trigger items specify a minimum p_T threshold for trigger objects. The thresholds can be set differently for each type of particle as well as for various combinations thereof. Trigger items may also require a coincidence of hits in the inner detector layers and/or the muon system. More complex trigger criteria exist, these will be discussed in their relevant sections below.

4.2.1 Level-1

The L1 trigger must cut the number of events by roughly a factor of 500 and has just 2.5 μs to process each event. The propagation of electrical signals from the detector to the L1 trigger system takes around a microsecond, leaving very little time for the trigger to make a decision. This tight time budget means the L1 trigger must be implemented in dedicated hardware.

The challenging timing conditions also limit the detector subsystems which can be used for the L1 trigger. The ID, CSC and MDT are unavailable for the L1 trigger. This is due to a combination of slow response and readout times, making them unsuitable for use at L1.

At L1, the RPC and TGC are used for triggering on high-p_T muons. As well as a minimum p_T requirement, the L1 muon trigger can also require a coincidence of hits in the muon system. This requirement improves the efficiency of muon candidates passing the subsequent trigger levels and offline identification cuts.

All calorimeter subsystems are available to the L1 trigger logic, albeit at a reduced granularity. This is typically 0.1×0.1 (0.2×0.2 for jets) in $\Delta\eta \times \Delta\phi$, with granularity decreasing at higher $|\eta|$. The calorimeters can be used to trigger on high-p_T electrons, photons, jets and hadronic taus at L1. Additional L1 triggers also exist which accept events with high amounts of missing transverse energy or total transverse energy. These are commonly referred to as E_T^{miss} and $\sum |E_T|$.

As well as cuts on p_T, the L1 calorimeter trigger is also capable of discarding events which fail simple isolation requirements. These configurable isolation requirements ensure energy deposits which are not sufficiently separated in angle from other calorimeter activity can be ignored. Cutting on isolation is available for

electrons, photons and hadronic taus at L1.

If a bunch crossing satisfies the muon or calorimeter trigger requirements described above, the event is accepted by the L1 trigger. The event is passed onto the next level of the trigger, for investigation in more detail.

4.2.2 Level-2

The L2 trigger is responsible for a reduction in the number of events by a factor of approximately 20. The design output for L2 is 3.5 kHz. L2 has significantly longer than L1 to make a decision, with an average time of 40 ms per event. Despite being much longer, this is still not sufficient time for all the data in every event to be read and used. The L2 trigger is implemented in software, making it more flexible but slower than L1.

L2 operates on the principle of using "regions of interest" (RoIs). These are cones of a fixed width, centred on the L1 trigger objects. The RoIs are read out from the detector and the data is passed to L2. Due to a limited area being read out for each event, the read out can be done at full granularity for all detector subsystems. Typically, one RoI accounts for 1–2% of the event data.

Again, L2 typically uses configurable p_T thresholds to determine whether an event should be accepted. L2 can also apply isolation and quality requirements to trigger objects. The inclusion of the ID allows electrons and photons to be distinguished more readily. Checks can also be made on muon candidates, to ensure their track is identifiable in both the ID and MS. Muons which satisfy this requirement are known as combined muons. L2 is also capable of running simple and fast selection algorithms.

4.2.3 Event Filter

The EF is the third and final trigger level, it must reduce the number of events by up to a factor of 20. This cuts the output stream to around 200 Hz, the rate at which the DAQ system can save events to disk for permanent storage. Similar to L2, the EF has access to full granularity output from all detector subsystems. Unlike L2 however, the EF has access to the entire event rather than just RoIs.

On average, the EF will have approximately four seconds to make a decision on each event. This allows much more detailed decision making compared to the lower levels of the trigger. Examples of this include the algorithms used to identify electrons and taus. These algorithms make a decision by inspecting variables, such as the ones which describe the particles shower shape in the calorimeters. Such

algorithms are discussed in more detail in Section 5.2.

If an event is accepted by the EF, a signal is passed to the DAQ requesting that the event be saved to disk. Depending on the trigger chain or chains which caused the event to be saved, the event is stored in one or more output streams. Example output streams include the "muon" and the "egamma" (electrons and photons) output streams.

4.3 Muon Trigger Scale Factors

A study of the 2011 muon trigger scale factors (SF) and their associated uncertainties has been performed. Trigger efficiencies describe how reliably a given object will cause a trigger to be activated. Scale factors are defined as the ratio between the data and MC efficiencies. They're applied to MC samples to ensure their behaviour matches that observed in data.

In this section, all quoted SF and uncertainties refer to those of the 2011 18 GeV muon trigger. The 2011 data is divided into lettered data periods, ranging from B to M. Each period is divided into numbered subperiods. For periods B–I, the trigger's name was `EF_mu18`. In the later data periods, muon quality requirements were required at trigger level and, accordingly, the trigger was renamed `EF_mu18_medium`.

The primary purpose of the study was to determine whether any SF uncertainties had been overlooked as well as to re-evaluate the previously calculated uncertainties. This was partially motivated by the requirements of the ATLAS precision measurement of the $W^{\pm}$ and Z^0 inclusive cross sections. The systematic uncertainty due to the muon trigger SF is one of the dominant uncertainties.

The use of a new simulated sample with improved statistics for the $Z \to \mu\mu$ process was one potential source for reductions in the uncertainties. If the statistical uncertainty due to the sample was partly or wholly responsible for the previously observed systematic uncertainties, then the new MC sample would cause them to be reduced. Replacing $\eta - \phi$ dependent scale factors with $\eta - p_T$ dependent ones was another potential source for improvements.

All efficiency and SF measurements described in this section are performed using the tag-and-probe method [36] applied to $Z \to \mu\mu$ events. $Z \to \mu\mu$ is ideal for measuring the muon trigger efficiencies and SF as it is a relatively common process which is reasonably easy to select with high purity. Events are required to have two oppositely charged, isolated muons with a combined invariant mass consistent with the Z^0 mass. It is also required that at least one of the muons passed the trigger that is being investigated, and this muon is labelled the tag muon. The efficiency

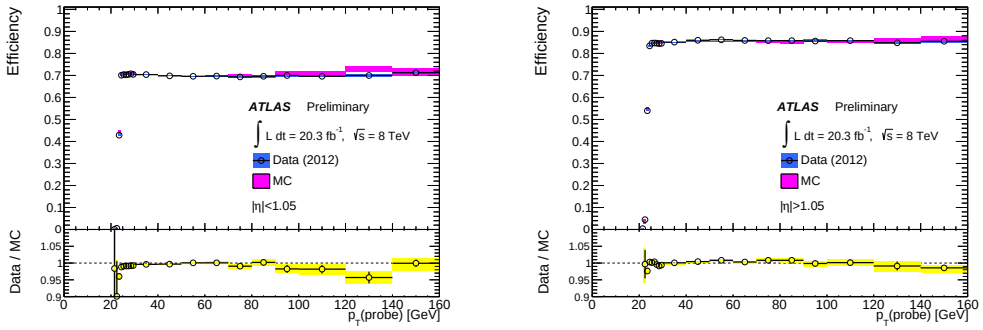

Figure 4.1: Muon trigger efficiency for the lowest p_T muon trigger in 2012 data, `mu24i_tight`, as a function of muon p_T [38]. Barrel (left) and endcap (right) muons are shown separately. A statistically significant variation with p_T can be seen.

may then be evaluated by examining whether the second muon in the event also activated the trigger.

4.3.1 Prior Status of the Muon Trigger Scale Factors

For the 2011 data, trigger SF and efficiencies are provided by the ATLAS muon trigger group binned in $\eta - \phi$ [37]. The combined uncertainty on the provided SF is $\pm 1\%$ and is made up of three components, combined in quadrature.

The SF exhibit some variation with p_T, as shown in Figure 4.1. However, the statistics of the available samples are too poor to bin the SF in p_T as well as $\eta - \phi$. A $\pm 0.7\%$ uncertainty is assigned to cover the SF fluctuations due to the transverse momentum of the muon.

The second systematic uncertainty is due to the choice of $\eta - \phi$ binning used for the SF. An area-weighted mean of the SF is calculated for the nominal binning, then compared with the mean for cases with finer and coarser binnings. A $\pm 0.6\%$ variation is observed and is taken as the uncertainty due to the binning choice.

The third uncertainty was assigned due to a potential bias caused by the preponderance of back-to-back tags and probes. The muon trigger's acceptance varies as a function of η and ϕ, due to the layout and properties of the muon trigger chambers. This means tags are more common in the high acceptance regions of the detector. The symmetry of the detector, combined with the fact that tags and probes are often produced back-to-back, means the majority of the probes are also found in high acceptance regions.

The sample is divided into two, depending on whether or not the tag and probe are found back-to-back. The exact cut for a tag and probe pair to be consid-

ered back-to-back is

$$\Delta\phi_{\mu\mu} > \pi - 0.1, \tag{4.1}$$

with events failing this cut forming the other subsample. The SF are found to vary $\pm 0.4\%$ between the two subsamples, therefore a third systematic uncertainty is assigned to account for this difference.

4.3.2 Study of the Muon Trigger Scale Factors

The primary aims of the study were as follows:

- Investigate additional potential uncertainties;
- Re-evaluate current uncertainties using new, high statistics MC sample.

The first additional uncertainty probed was the uncertainty introduced by the background estimation technique used when calculating the SF. Here, background processes are any non-$Z \to \mu\mu$ events which are present in the data samples used to calculate the SF. The backgrounds included in the study were Z+jets (with the Z decaying to either ee or $\tau\tau$), W+jets, top-quark related (both $t\bar{t}$ and single top processes) and diboson (WW, ZZ and WZ).

Two background treatment methods were compared in the study. For the control case, each background sample was normalised to its expected number of events using its cross section. Each background contribution was then subtracted from the data before the SF were calculated.

For the first test case, the normalisation of each background was varied up and down by the 1σ uncertainty on its cross section. The SF were then recalculated with the varied backgrounds subtracted and the change in SF as compared to the nominal case was taken as the uncertainty.

For a second, more extreme, variation, the SF were recalculated treating the data sample as pure $Z \to \mu\mu$. This means that no background subtraction was undertaken before the SF were determined. Again, the SF variation between this and the nominal case was taken as the uncertainty for this scenario.

The variation in SF for both cases was found to be orders of magnitudes smaller than the SF's statistical uncertainty. This is primarily due to the applied $Z \to \mu\mu$ selection being highly pure, with minimal background contamination. For these reasons, no additional uncertainty was deemed necessary. The results are summarised in Figure 4.2.

The SF were originally derived for the charge inclusive case. A comparison was made between SF derived separately for positively and negatively charged

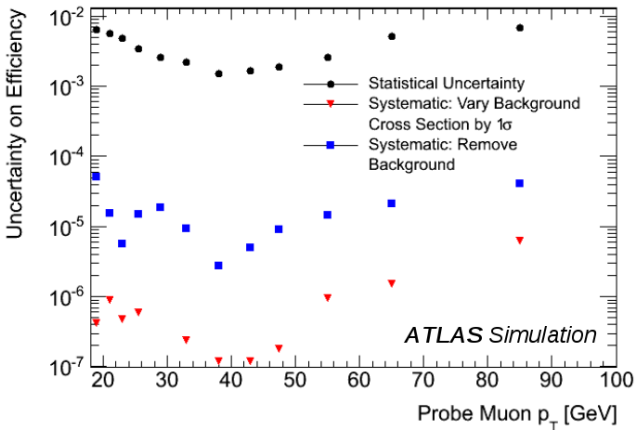

Figure 4.2: Results of the study of how the background model treatment impacts the muon trigger efficiency calculations. The black points show the size of the statistical uncertainty on the efficiencies. The red points show the uncertainty due to varying the cross sections of the background processes within their uncertainties. The blue points show how much the muon trigger efficiencies change if the background processes are simply ignored. It can be seen that the potential background modelling uncertainties are orders of magnitude smaller than the statistical uncertainty.

muons. No significant differences were observed for the late-2011 data periods (G6-M). However, significant variation was seen in regions of the detector when comparing positively and negatively charged muons for the data periods B-G5. The charge asymmetry can be seen in Figure 4.3.

The discrepancy was associated with a bug in a muon trigger algorithm. This algorithm was responsible for associating ID and MS tracks for the Level-2 trigger. The bug was fixed following period G5, hence the disappearance of the effect in the later data periods. However, it was decided that any new 2011 SF would be calculated separately for positively and negatively charged muons. This forgoes the need for an extra systematic uncertainty.

The first systematic to be re-evaluated was the back-to-back bias uncertainty. When using the $Z \to \mu\mu$ sample with improved statistics, no significant differences were observed between SF derived using the two subsamples (of events where the tag and probes are and are not back-to-back). The conclusion was that the previously calculated systematic uncertainty was merely due to statistical fluctuations in the old MC sample, and therefore could be safely discarded.

Similarly, the binning systematic was also found to be partially due to statistical fluctuations. Comparisons between the area-weighted mean SF for the nominal, coarse and fine binning configurations suggested the systematic uncertainty was of the order $\pm 0.2\%$, rather than the originally quoted $\pm 0.6\%$. This reduced systematic

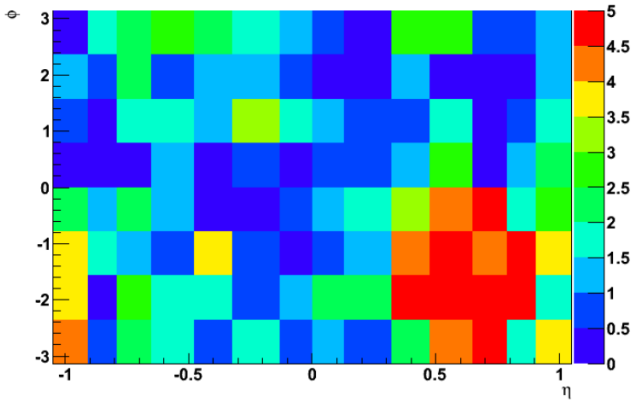

Figure 4.3: A visualisation of the charge asymmetry of the muon trigger efficiency for data periods B-G5. The plot shows the significance of the charge asymmetry in bins of $\eta - \phi$, defined as $|\epsilon_+ - \epsilon_-|/\sigma_\epsilon$. The significant charge asymmetry can be observed in the $(+\eta, -\phi)$ quadrant.

uncertainty was recommended for use in the future.

The uncertainty due to the SF variation with p_T did not prove as easy to reduce. Attempts were made to identify a relationship between p_T and SF but no consistent relationship was found. Furthermore, re-calculating the uncertainty with the higher statistics MC sample did not result in a reduced systematic uncertainty. The previously mentioned requirements of the $W^{\pm}/Z^0$ precision analysis meant a different approach had to be made.

The decision was made to calculate SF using the same binning as the $W^{\pm}/Z^0$ analysis. The analysis uses an $\eta - p_T$ binning, rather than binning in $\eta - \phi$. The new SF were derived separately for positively and negatively charged muons, due to the significant differences observed in the aforementioned study.

Ideally, the SF would be calculated binned in η, ϕ and p_T but the statistics were insufficient to bin the SF calculation in three variables simultaneously. Physics processes can exhibit a dependence on p_T and η, meaning SF variations in these two variables are most important to parameterise accurately. This is the reason behind choosing $\eta - p_T$ as the primary SF binning.

Following the derivation and application of the new $\eta - p_T$ SF, "second-order" $\eta - \phi$ SF were derived. These second-order SF should reduce the effect of any residual variation of the SF as a function of ϕ.

Following these calculations, the methodology and results were presented to and approved by the ATLAS muon trigger group. The SF were then utilised by the $W^{\pm}/Z^0$ analysis. The analysis is currently awaiting publication.

The scale factors are given in Appendix A. They are provided separately for positively and negatively charged muons and divided into four macro-periods. Periods B–G5 represent the data period where a large difference in SF was seen for positively and negatively charged muons for the `EF_mu18` trigger. This was corrected between periods G5 and G6. The second macro-period, G6–I, represents the remainder of data recorded using the `EF_mu18` trigger. The vast majority of data taken using the `EF_mu18_medium` trigger makes up the third macro-period, periods J–M (excluding L3 and L4). Periods L3 and L4 form the fourth macro-period, due to a bug in the trigger code resulting in vastly different SF when compared to the rest of periods J–M.

Chapter 5

Selection and Categorisation

5.1 Introduction

This thesis describes a search for the Higgs decaying to a pair of tau leptons, where one tau decays hadronically and the other decays to a light lepton. This chapter will cover a number of key steps in the $H \to \tau_\ell \tau_{had}$ analysis. It will begin by discussing event reconstruction, where the detector readout is interpreted as the physics objects used in the analysis. It will then describe how the events of the analysis are selected and categorised.

5.2 Event Reconstruction

Event reconstruction is an important process which all events must pass through before analysis can begin. Whether data recorded by the ATLAS detector, or simulated MC samples, all events pass through the same ATLAS reconstruction software. Here, the techniques used to convert tracker hits and energy deposits to analysis level objects will be discussed. All object categories relevant to this analysis will be covered in turn.

5.2.1 Muons

In ATLAS, muons are constructed using hits in the inner detector and muon spectrometer. Some limited information from the calorimeters is also used. During the event reconstruction process, muon candidate tracks are constructed separately in the ID and MS.

Reconstruction

The MS is capable of precisely measuring muon tracks using the MDT and CSC. In the bending plane, each subdetector layer has a resolution in the range 60–80 μm. First, a search is made for track segments in the individual layers of the MS. These segments are then combined across multiple layers to form a full MS track.

Tracks in the ID are reconstructed in a similar manner, with additional track quality requirements [39]. Each muon track is required to have at least one hit in the pixel detector and five or more in the SCT. A track which crosses more than two layers with no hits recorded is also discarded. These requirements are relaxed if a sensor along the track has been flagged as inefficient. In the region of TRT coverage, $0.1 < |\eta| < 1.9$, nine or more TRT hits are also required.

ATLAS defines four classes of muons, each with different specific requirements. These are:

- Standalone (SA) muons: These muons are defined using candidate tracks from the MS only. The tracks are extrapolated back towards the beamline, with a correction made for energy lost to the calorimeters. The interaction point is defined as the point of closest approach between the beamline and extrapolated track. The other muon classes rely on the ID, which is only instrumented for $|\eta| < 2.5$. SA muons extend coverage over the range $2.5 < |\eta| < 2.7$.

- Combined (CB) muons: These are muons where tracks are independently reconstructed in the ID and MS. A combined muon is then formed in cases where an ID and MS track can be successfully joined. CB muons require the MS track to have hits in multiple layers of the MS. Due to the tight quality constraints on the two independent tracks, CB muons offer the best purity.

- Segment tagged (ST) muons: Similar to CB muons except only one hit is required in the MS. This recovers some efficiency for muon candidates with $p_T < 10$ GeV or muons which passed the MS in a region with reduced acceptance. Almost 20% efficiency is recovered at $p_T^{\mu} = 5$ GeV and up to 35% is gained for muons with $|\eta| \approx 1.2$, as shown in Figure 5.1.

- Calorimeter tagged (CaloTag) muons: The ID candidate track is combined with calorimeter deposits consistent with a minimum ionising particle (MIP). These muons recover some acceptance in regions where the MS is uninstrumented.

Comparisons of the performance of the first three algorithms as a function of muon p_T and $|\eta|$ are shown in Figure 5.1.

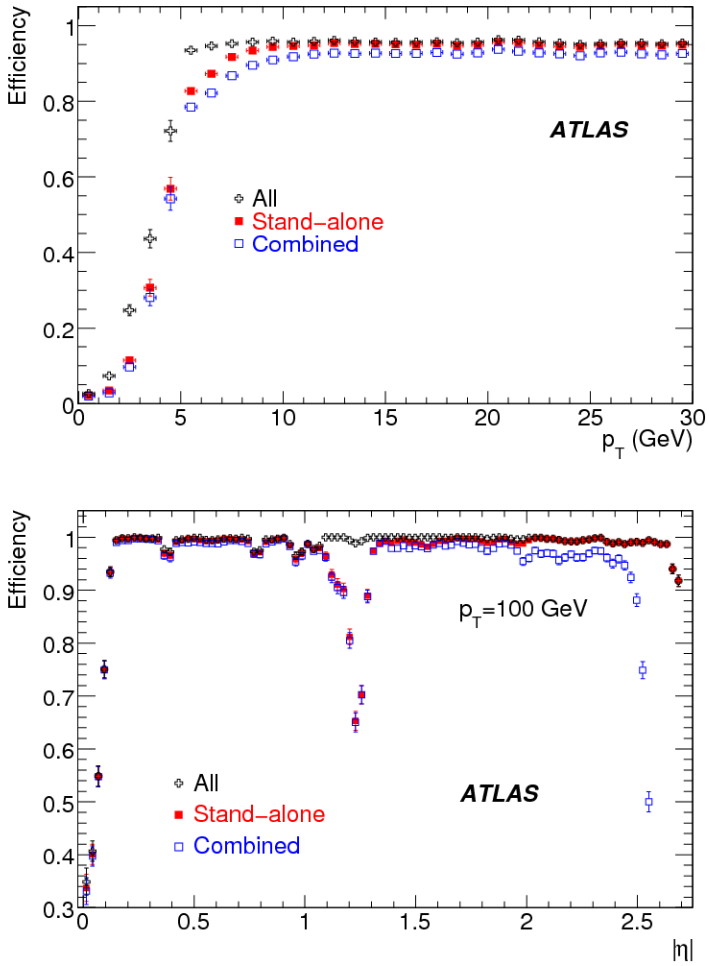

Figure 5.1: Muon reconstruction efficiency as a function of p_T (top) and $|\eta|$ (bottom) [29]. The efficiency is shown for standalone muons (red), combined muons (blue) and for the combination of standalone, combined and segment tagged muons (black).

There are two noteworthy regions where the MS lacks in coverage. The first is the region $|\eta| < 0.1$, where there is a crack between the two halves of the MS barrel. The second is a gap between the barrel and positive-eta endcap, $1.1 < \eta < 1.9$. Here, not all of the planned MS chambers were installed. These two regions can clearly be seen in Figure 5.1.

Performance

Muon trigger efficiencies are calculated using a tag-and-probe method (see Section 4.3 for an explanation of trigger efficiencies and the tag-and-probe method). For muons with $p_T > 25$ GeV, the overall muon trigger efficiency is around 70% in the barrel and 85% in the endcaps [38]. The L1 trigger is responsible for the vast majority of the efficiency loss.

The reconstruction efficiency is generally around 99% for both CB and ST muons [39], with the two regions listed above being the major exceptions. The crack region is roughly 66% efficient for CB and ST muons. The positive-eta gap has an efficiency of approximately 85% for CB muons but this climbs to 99% when considering CB and ST muons.

The trigger and offline reconstruction efficiencies are calculated in both data and MC simulation, with no major mismodellings observed. Separate trigger and reconstruction efficiency corrections are derived for MC to correct for any minor discrepancies.

Another important measure of performance is the p_T resolution. This is the uncertainty on the momentum measurement, typically expressed as a percentage of the total p_T of a particle. For the CB muons used in this analysis, the momentum resolution is of the order 2% [39].

Selection

This analysis considers only CB and ST muons. Cuts are also applied, selecting muons with $|\eta| < 2.5$ and $p_T^\mu > 10$ GeV.

5.2.2 Electrons

The reconstruction and identification of electrons depends predominantly on the inner detector and electromagnetic calorimeter. Making use of the hadronic calorimeter makes the rejection of hadronic jets more straightforward. An ECAL energy deposit matched with a track is treated as an electron candidate. Additional details on the reconstruction process can be found below.

The TRT and the first layer of the pixel detector are especially useful for electron identification. X-ray transition radiation produced by electrons in the TRT allow them to be more readily distinguished from charged hadrons. The first pixel layer makes identifying electrons produced by photon conversions easier [40]. This is because photon conversions will typically not leave a hit in the first layer.

Reconstruction

The first stage of electron reconstruction is the identification of an ECAL energy cluster. The ECAL is treated as a 200×256 grid in $\eta \times \phi$. Each tower in the grid has $\Delta\eta \times \Delta\phi$ dimensions of 0.025×0.025. The energy is summed across the three ECAL layers, including the presampler where relevant, for each tower. ECAL cells which overlap multiple towers have their energy equally divided amongst them.

These towers are then scanned over, with any 3×5 group of towers with $E_T > 2.5$ GeV being flagged as an ECAL cluster. Methods are in place to identify and remove nearby duplicate clusters. The algorithm is named the sliding window algorithm [41]. For a true electron with a p_T of 15 GeV, the cluster finding algorithm is around 99% efficient in the barrel [40]. The final step in cluster finding is discarding clusters with significant energy deposits in the HCAL or smeared too widely in η. Such clusters are found to be inconsistent with true electron shower shapes.

The next stage is the identification of electron-track candidates. An RoI is defined about the centre of each ECAL cluster, with $\Delta R = 0.3$. Any track seed with three hits in the silicon detector and $p_T > 1$ GeV is identified. Then, tracks which are consistent with a $\pi^{\pm}$ are discarded. Next, tracks which are located in the cluster RoI are fully reconstructed, using an electron-like model for energy loss at material boundaries.

These ID tracks are then extrapolated to the ECAL. If the track is found to be consistent with the ECAL energy cluster centre, a more detailed track refitting is performed. This refitting corrects for electron energy losses due to transition radiation and bremsstrahlung.

At this stage in the reconstruction process, each cluster may be associated with multiple track candidates. The best matched track is now selected, although all matched tracks are retained. The track selection prefers well matched tracks with multiple hits in the pixel tracker. The choice of track is key as the best matched track is used to calculate the electron charge and kinematics.The cluster and track combination forms the electron candidate, ready for inspection by the electron identification algorithm.

Identification

The additional step of electron identification is required as the electron detector signature is not as clean as that of muons. This means that a significant proportion of candidate electrons are not isolated true electrons, which are of most interest in analyses. In this context "identification" takes on the meaning of additional cuts which efficiently select isolated true (signal) electrons while rejecting non-isolated and fake (background) electrons. Electrons can be faked by a number of sources, including photon conversions and hadronic jets. Non-isolated electrons arise from processes such as semi-leptonic decays of heavy flavour hadrons. The identification procedure is designed to be able to discriminate against these processes.

This analysis uses the cut-based electron discriminant. It cuts on a number of variables relating to the properties of the track, the cluster shower shape and the matching between the track and cluster. The variables which are cut on are described in Table 5.1. This variable list corresponds to the `medium` working point defined by the electron group [42]. The electron candidates are divided into E_T and η bins, with the cut values individually optimised in each bin. This optimisation of the cut values gives the best signal efficiency for a given background rejection.

Performance

The performance of the electron reconstruction and identification algorithms has been carefully studied [40, 42]. The relevant measurements will be briefly summarised here. The efficiencies vary as a function of year, the number of pile-up vertices and the electron E_T and η. They are measured separately for data and MC simulation, with scale factors derived to correct any mismodellings.

The reconstruction efficiency for electrons with $p_T > 20$ GeV varies between 90 and 99%, with the barrel reconstruction efficiency outperforming that in the end-caps. The efficiency in 2012 is typically 5% better than in 2011 due to improvements in bremsstrahlung modelling when fitting the electron track. For 2011 samples, the identification efficiency for the `medium` working point is in the range 75–90%. The figure is again higher for 2012, spanning 80–99%, due to a reoptimisation of the cuts. The MC correction factors are typically around 2%, increasing at low E_T and high-η.

The electron E_T resolution has also been carefully measured using $Z \to ee$ events [43]. The resolution varies with electron E_T and η but is generally of the order of 1%.

Cut Category	Cut Description
Electron acceptance	$\|\eta\| < 2.47$
Hadronic leakage	Ratio of E_T deposited in the hadronic calorimeter to E_T of the EM cluster
Middle layer of EM calorimeter	Ratio of total energy in 3×7 cluster of calorimeter towers, centred on electron candidate, to that in 7×7 cluster Lateral shower width
Front layer of EM calorimeter	Total shower width Difference between two largest energy deposits, divided by their sum
Track quality and cluster matching	Number of hits in the b-layer (first pixel layer) Number of hits in the pixel detector Number of hits in the silicon detector $\|\Delta\eta\|$ between extrapolated track and cluster in first layer of EM calorimeter Transverse impact parameter, d_0, the distance of closest approach between track and beam axis
TRT	Fraction of TRT straws activated at high-threshold

Table 5.1: Variables used in the cut-based electron discriminant [42]. The cut values are optimised to provide the biggest signal to background separation in bins of E_T and η.

Selection

As well as requiring electrons to pass the `medium` electron identification cuts, the analysis requires a number of other electron selection cuts. Electrons must have a minimum p_T of 15 GeV and $|\eta| < 2.47$. Furthermore, electrons with $1.37 < |\eta| < 1.52$ are discarded. This corresponds to the gaps between the ECAL barrel and endcaps. Here there is an excess of material between the interaction point and calorimeters, degrading the performance of electron identification and measurement.

5.2.3 Jets

Hadronic jets are reconstructed solely using information taken from the hadronic calorimeter of ATLAS. Jets are reconstructed using the anti-k_T algorithm [44], seeded by topological calorimeter clusters [41].

Cluster Finding

Firstly, the process of forming the topological clusters will be briefly described [41]. The first stage of topological clustering uses the signal-to-noise ratio (SNR) of each cell as the clustering variable. The SNR provides a measure of the significance of any energy deposit in the calorimeter. The noise term consists of both electronic noise and a noise term relating to the pile-up.

A list of all calorimeter cells with SNR > 4 is produced as a starting point. The high SNR threshold requirement is successful at suppressing contributions from noise and pile-up. Starting with the highest SNR seeds, each seed is iterated through with all neighbouring cells added to form the initial clusters. Neighbouring cells in this context refers to the eight surrounding cells in the same layer, as well as the overlapping cells in adjacent layers.

If a neighbour cell has an SNR > 2, it is added to the list of seeds for the next iteration. If such a neighbour cell borders two clusters then the two are merged. At this stage, all initial seed cells have absorbed their neighbouring cells into their cluster. Clusters which share high-SNR neighbours have been merged and all high-SNR neighbours have been added to a new list of seeds.

This process is iterated, using this new list as the seeds for the next iteration. It continues until an iteration where the list of new high-SNR neighbours is empty, leaving no seeds for the next pass through.

The algorithm now enters the second stage of topological clustering. This stage uses the cell energies to re-form the clusters from the previous stage. It is seeded with a list of local maxima. These seed cells are required to have $E > 500$

MeV, no neighbouring cells with a higher energy and at least four neighbours in the same stage-one cluster.

Similar to the previous stage, each seed cell's neighbours are iteratively added to the seed's cluster. However, only cells used in the previous stage are considered. The process ends when all cells have been assigned to a cluster. Any cells adjacent to two clusters are split between the two. Any stage-one cluster without a cell satisfying the local maxima criteria listed above is left unchanged.

These final clusters are the output of the topological clustering algorithm. The energy of each cluster is calculated by summing the cell energies. These clusters are then sorted by energy before being passed as input to the anti-k_T jet clustering algorithm.

Jet Clustering

The standard jet clustering algorithm used by ATLAS is the anti-k_T algorithm [44]. It is an evolution of the k_T jet clustering algorithm, designed to tackle its shortcomings.

Two key features of a successful jet clustering algorithm are that it must be infrared and collinear safe. Infrared safe means that soft emissions should not change the output of the algorithm. Collinear safe means that the jet should be clustered identically, no matter how collinear contributions are divided. For example, a single object should be clustered similarly to two collinear objects, each with half the energy.

The k_T family of algorithms [45] work by combining clusters into jets, starting with the two objects separated by the smallest distance, d_{ij}. This distance metric is defined as

$$d_{ij} = \min(p_{Ti}^{2p}, p_{Tj}^{2p}) \frac{\Delta R_{ij}^2}{R^2}. \tag{5.1}$$

However, if the distance between an object and the beampipe, d_{iB}, is less than the smallest d_{ij}, the object is saved as a jet and removed from further consideration. We define

$$d_{iB} = p_{Ti}^{2p}. \tag{5.2}$$

The p_T and ΔR variables have been described previously. R is a parameter loosely related to the radius of the outputted jets. This analysis uses $R = 0.4$ to cluster jets. The parameter p can be adjusted, giving different styles of algorithms for negative and positive values.

The k_T algorithm uses $p = 1$. By inspecting the distance metric defined above, we can see how the algorithm behaves. It works by clustering nearby low p_T

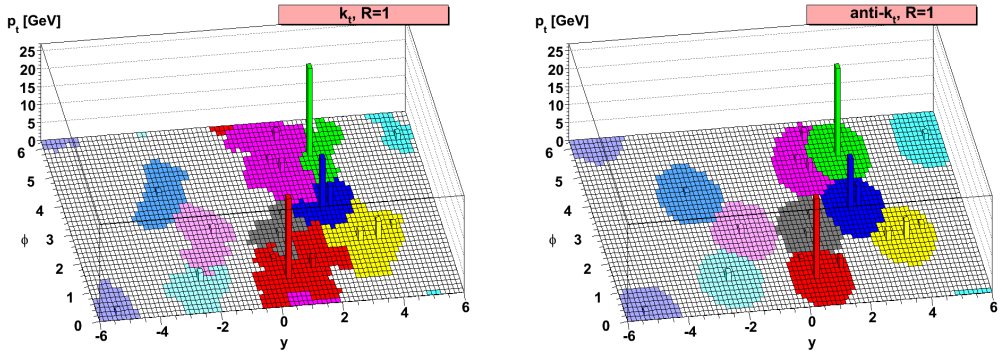

Figure 5.2: Example of the behaviour of the k_T and anti-k_T jet clustering algorithms [44]. The two algorithms use identical inputs, making direct comparisons possible.

objects, continuing until all input clusters have been added to a jet. While being infrared and collinear safe, the k_T algorithm outputs irregularly shaped jets.

The anti-k_T algorithm derives its name from the fact it uses $p = -1$. The change has the effect that objects cluster to nearby ($\Delta R < R$) high-p_T objects. This means the highest p_T objects form the seeds of perfectly conical clusters.

If two nearby jet seeds ($\Delta R < 2R$) have similar transverse momenta, they form approximate cones with objects between them allocated to the nearest jet. This gives two overlapping cones with a straight boundary between them. Intermediate p_T objects near higher p_T ones form cones with a segment lost to the higher p_T seed. The anti-k_T algorithm is infrared and collinear safe, with the added feature that the jet boundaries are not affected by nearby soft radiation.

Figure 5.2 shows an example of jet clustering for both the k_T and anti-k_T jet clustering algorithms. The various features described above are clearly visible.

b-jet Vetoes

ATLAS has developed algorithms which allow jets originating from b-quarks to be identified. Identifying and vetoing events containing b-jets is useful for reducing the background contribution from top events. This process is commonly referred to as b-tagging.

The primary algorithm used by ATLAS is the MV1 algorithm [46]. It takes as input the output from other b-tagging algorithms, namely IP3D, SV1 and JetFitter. These algorithms discriminate between regular and b-jets using a range of variables.

Due to their non-negligible lifetimes, b-quark resonance decays (secondary vertices) are slightly displaced from the primary vertex of an event. IP3D uses variables relating to the impact parameter, the distance of closest approach between

a particle track and the beam axis. Similarly, SV1 looks for a secondary vertex characteristic of a b-quark decay. JetFitter, on the other hand, studies the topology of the hadronic decay in an effort to differentiate between b and non-b decays.

This analysis uses the MV1 working point where the algorithm is on average 70% efficient at identifying jets originating from b-quarks with $p_T > 15$ GeV. Depending on p_T and η, the misclassification rate varies between 0.1 and 0.5% [46]. Scale factors are derived to correct for any mismodellings in simulation, with a typical size of 2–4%.

Performance

One key measure of jet performance is the jet energy resolution. Low p_T jets have a typical resolution of 15%, decreasing to around 5% as jet p_T increases [47]. Some dependence on η is also observed. The jet energy resolution is observed to be well modelled in MC simulation.

The jet energy scale is the second factor which is important to the jet performance. It quantifies the relationship between the detector read out and the energies of reconstructed jets, varying as a function of the detector region. An accurate calibration of the jet energy scale is vital for the proper measurement of jet energy. The calibration is carried out using dijet and multijet events in data and MC [47].

Selection

Only jets with $|\eta| < 4.5$ are considered in the analysis. The η requirement is looser than that of other objects (e, μ and τ) as jet reconstruction does not rely on the ID.

Jets with $|\eta| < 2.5$ are evaluated by the b-tagging algorithm. Any event with a b-tagged jet is discarded, significantly reducing the number of top related background events.

One final requirement is applied to jets with $p_T < 50$ GeV and $|\eta| < 2.4$, regarding the jet vertex fraction (JVF) [48]. The JVF for a jet is defined as the sum of the transverse momenta of tracks in the jet assigned to the primary vertex, divided by the total sum transverse momentum of all tracks in the jet. For 7 (8) TeV events, the requirement is JVF > 0.75 (0.5). The cut was loosened for 8 TeV events due to the increase in pile-up.

5.2.4 Hadronic Taus

Hadronically decaying taus are primarily reconstructed using the hadronic calorimeter. Additionally, the ID provides important information on the number of charged

tracks in the tau decay. Taus decay hadronically roughly 65% of the time, with the remaining 35% being leptonic decays [1]. The majority (nearly 60% of the total) of the hadronic decays include one or three charged pions and zero to three neutral pions. The most common hadronic decay channels are $\pi^-\pi^0\nu_\tau$ (25.5%), $\pi^-\nu_\tau$ (10.8%), $\pi^-\pi^0\pi^0\nu_\tau$ (9.3%) and $\pi^-\pi^+\pi^-\nu_\tau$ (9.0%). The charged pions leave tracks in the ID which are important for tau identification.

Reconstruction

The tau reconstruction and identification is seeded by the jets described in the previous section [49]. The parameter R is kept as 0.4 but only jets with $p_T > 10$ GeV and $|\eta| < 2.5$ are considered. Tau reconstruction also requires events to have a primary vertex with three or more associated tracks.

Each jet is split into two regions when considered as a tau candidate. The first is known as the core region and is defined as the volume where $R < 0.2$. The second is the isolation region and covers the range $0.2 < R < 0.4$.

To be considered a tau candidate track, a track must

- be in the core region
- have $p_T > 1$ GeV
- have at least two pixel hits
- have five or more additional silicon hits
- have $|d_0| < 1.0$ mm
- have $|z_0 \sin\theta| < 1.5$ mm

Tau candidates require exactly one or three such tracks to be considered. Tracks which satisfy the above conditions but are located in the isolation region are also kept for use in the tau identification.

Before running the tau identification algorithm, effort is made to reconstruct neutral pions, π^0, in the core region. Each core region may contain between zero and two π^0. Neutral pions predominantly decay to two photons, meaning each π^0 candidate is observed as two energy clusters. Energy from pile-up, the underlying event and calorimetric noise is subtracted. The properties of the energy clusters are used to calculate how likely the π^0 candidate is to be a true π^0.

The ID tracks passing the requirements listed above and the highest scoring π^0 candidates are passed to the tau identification software.

Identification

The primary function of the tau identification machinery is to discriminate between hadronic taus and jets. As taus are seeded by jets, the identification step is important to ensure taus and jets are not misidentified as each other. As well as the jet discriminant, there is also an electron veto to identify and discard electrons misidentified as taus.

The jet discriminant software uses a boosted decision tree (BDT) to classify the tau candidates. Separate BDTs are trained for one- and three-track tau candidates. The BDTs use a number of variables which will be listed here. A full explanation of the variables can be found elsewhere [49].

The discriminant variables relate to the

- energy distributions in the calorimeters
- proportion of the energy in the lead track
- track directions
- number of tracks in the isolation region
- secondary vertex properties
- invariant mass of the tracks, neutral pions and the ratio between them
- number of π^0.

Different lists of the variables are used by the one- and three-track BDTs.

The BDT has three working points corresponding to different signal efficiencies. This analysis uses the `medium` working point, corresponding to a 60% (40%) efficiency for one (three) track taus. The main sources of inefficiency are due to track reconstruction inefficiencies and additional tracks being included from the underlying event.

The electron veto is also BDT-based and is designed to identify and discard electrons faking taus. Variables used by the electron veto include the transition radiation properties, angle of the track, deposits in the electromagnetic calorimeter and the calorimeter shower shape. This analysis uses the `loose` electron veto which has a efficiency for true taus of 95%. No dedicated muon veto is used in the analysis due to their negligible expected contribution. This is because muons are minimum ionising particles and therefore deposit very little energy in the calorimeters, greatly lowering the chance of misidentification.

Performance

The signal efficiency and background rejection of the tau identification machinery are important measures of performance. The cut values on the BDT jet discriminant vary as a function of tau candidate p_T, chosen to maintain a constant signal efficiency across the different p_T bins. For one track taus, the signal efficiency of the `medium` working point is set to 60% [49]. This corresponds to a 95% reduction in the number of jets faking taus. The chosen signal efficiency for three track taus is 40%. This gives a considerably better background rejection with only 0.2–1% of jets passing the BDT cut, varying with the p_T of the tau candidate.

The second important performance consideration is the effectiveness of the electron veto. The electron veto is only available for one track taus as these are much more likely to be faked by electrons. The `loose` working point is used, giving a 95% signal efficiency. The number of electrons faking taus is substantially reduced by the application of the veto. Varying as a function of η, 95–97.5% of electrons are rejected.

The tau energy resolution of the ATLAS detector has also been measured [49]. For taus with $E_T = 20$ GeV, the resolution is typically around 20%. There is, however, some variation depending on the tau candidate's number of tracks and η. The resolution decreases as energy increases, to of the order 10% for taus with $E_T = 100$ GeV.

Selection

The primary kinematic tau selections used in the analysis are $p_T > 20$ GeV, $|\eta_{\text{lead\_track}}| < 2.47$, a charge of ± 1 and one or three tracks in the core region ($R < 0.2$). In addition, taus must pass the `medium` selection of the tau-jet discriminant and the `loose` electron veto.

5.2.5 Object Overlap Removal

In ATLAS it is possible for detector deposits to be reconstructed as multiple objects. For example, given specific circumstances, a calorimeter deposit can be reconstructed as an electron, a hadronically decaying tau and a jet. Object overlap removal (OOR) is the analysis step where such overlapping objects are considered, giving priority to objects in a specific order and discarding the rest.

In this analysis, objects undergo OOR when separated in ΔR by less than 0.2. Muons are given the highest priority (hence are never discarded by OOR), followed by electrons, hadronic taus and finally jets.

For the OOR, some object definitions are relaxed. This has been carefully studied and found to significantly reduce occurrences of objects faking other objects. The muon p_T threshold is reduced to 2 GeV for OOR against taus and the track quality requirements are also relaxed slightly. Unlike in the analysis, electrons in the crack region ($1.37 < |\eta| < 1.52$) are considered for OOR. Additionally, the quality requirements for non-crack electrons is reduced from `medium` to `loose` for OOR. These relaxations of the electron definition mean tau and jet candidates which are actually poorly reconstructed electrons are correctly discarded.

5.2.6 Missing Transverse Energy

As the colliding protons of both bunches travel parallel to the z-axis, the transverse energy of any interaction should sum to zero. For a number of reasons particles are able to escape detection and the resulting energy imbalance is named the missing transverse energy, E_T^{miss}. Some particles simply pass through gaps in the detector coverage, most notably through the gaps between the barrel and endcap detector components. Also, particles such as neutrinos are so weakly interacting that they are able to pass through the entire detector without interacting with the instrumentation.

The E_T^{miss} is primarily calculated using information from the calorimeters. They absorb the bulk of the energy of interacting particles and provide good coverage up to high-η. In addition, information on muons and low-p_T particles is taken from the MS and ID respectively.

In the $H \to \tau_\ell \tau_{had}$ analysis, each signal event has three neutrinos. Both tau decays produce a neutrino with the third produced when one of the secondary W bosons decays to a light lepton. This means the E_T^{miss} reconstruction performance has a significant impact on the analysis.

Reconstruction

The E_T^{miss} for each event is calculated using fully reconstructed objects [50]. Energy deposits in the calorimeters are associated to these objects in a specified order. The highest priority is given to electrons, followed by photons, hadronic taus, jets and finally muons. The main energy calculation for muons however is performed using the dedicated muon system.

An additional soft term is calculated which includes all calorimeter energy deposits not associated with an object. In an effort to suppress calorimeter noise, the clusters used are topological clusters [41]. This soft term also encompasses the

low-p_T ID tracks with no associated objects.

A transverse energy vector in the $x-y$ plane is defined for each object in the event, taking the object's transverse energy as the vector's magnitude. The E_T^{miss} is simply calculated by taking the negative of the vector sum of these transverse energy vectors. This gives the vector which corresponds to the total missing transverse energy in the event.

Pile-up has a significant impact on the E_T^{miss} calculation so effort is made to suppress its effect. This analysis uses the Soft Term Vertex Fraction (STVF) method [50]. The p_T fraction of soft term tracks originating from the primary vertex is calculated. Soft term tracks simply means ID tracks not associated to any physics objects. The soft term in the E_T^{miss} summation is then scaled by this factor, reducing the impact from pile-up.

5.2.7 Higgs Boson

Although this analysis is a search for the Higgs boson, only its decay products are observed in the detector. This means it is important to be able to reconstruct the candidate Higgs in each event from the measured objects. One reason why this is important is that the mass of the parent particle is an important discriminating variable. While $H \to \tau\tau$ and $Z \to \tau\tau$ both produce similar signatures, the difference in mass between the Higgs and Z bosons allows some distinction to be made.

Due to the presence of three neutrinos in the final state, reconstructing the Higgs is complex and may be imprecise. The E_T^{miss} performance is of particular importance as this is the primary handle on the three neutrinos. ATLAS use a tool named the Missing Mass Calculator (MMC) [51] to reconstruct the Higgs candidate from the measured decay products.

In a $H \to \tau_\ell\tau_{had}$ event, there are seven unknowns in total. The x-, y- and z-components of momentum for the two tau neutrinos form the first six. The neutrino which originates from the leptonic tau decay adds only one additional unknown, its invariant mass. Conservation of energy and momentum define the tau neutrino masses and the lepton neutrino momentum.

The seven unknowns are further constrained by four relationships. The x- and y-components of the E_T^{miss} must equal the sum of the x- and y-components of the neutrino momenta. Furthermore, the mass of a tau lepton is known, constraining the total invariant mass of each tau decay system. However, seven unknowns with only four constraints leaves the situation under-constrained for directly solving.

The MMC takes the constraints and scans over ranges of values for the unknown quantities. Not all combinations have an equal chance of occurring, therefore

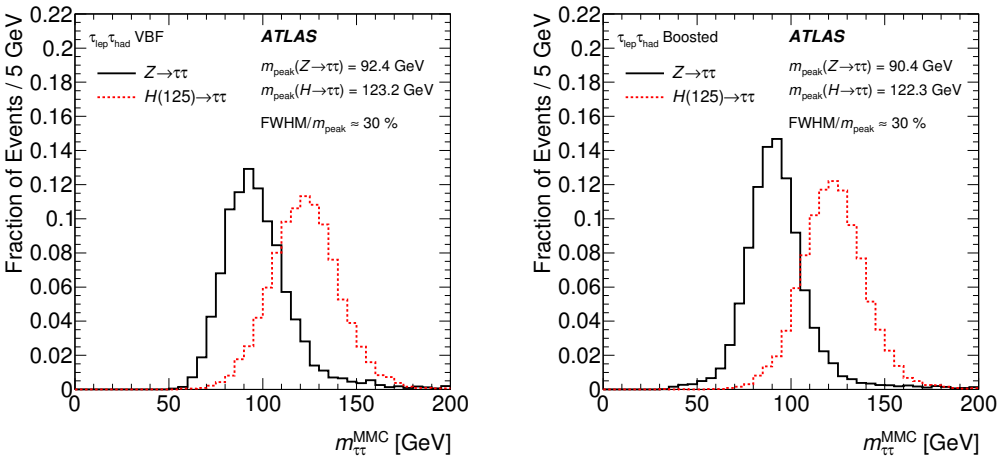

Figure 5.3: Comparison of the ditau invariant mass, as calculated by the Missing Mass Calculator (MMC) [51], for $H \to \tau\tau$ and $Z \to \tau\tau$ in the different analysis categories [52].

the likelihood of each scenario is calculated alongside the resulting ditau mass. An example of a variable which contributes to the likelihood calculation is the ΔR between the visible and invisible decay products of each tau. Its expected distribution is well known and this is converted to a probability density function and included in the likelihood.

Finally, the ditau mass with the greatest total likelihood is selected for use. The combined process is 99% efficient for signal events and the mass resolution is around 16%. Histograms comparing the mass reconstruction for $H \to \tau\tau$ and $Z \to \tau\tau$ can be seen in Figure 5.3.

Another useful variable relating to the Higgs candidate is its transverse momentum, p_T^H. It is defined as the vector sum of the transverse momenta of the visible decay products and the E_T^{miss}. One of its uses is in identifying events where the Higgs boson is heavily boosted.

5.3 Event Selection

Before categorising events, a preselection is made to select good quality events with the desired collection of objects. Some object definitions are tightened before the event selection cuts are made.

5.3.1 Triggers

Selecting the triggers to use for an analysis is one of the first steps when selecting events. The analysis uses low-p_T single lepton triggers. These triggers have a high efficiency for $H \to \tau_\ell \tau_{had}$ events and are simpler than triggers with a tau requirement.

For the start of 2011 data taking, the electron trigger threshold was 20 GeV and the muon one was 18 GeV. As the luminosity increased throughout the year, the trigger requirements were tightened to keep their rates manageable. The electron trigger threshold increased to 22 GeV and the tracking requirements were tightened for both electrons and muons. For 2012 data taking, the electron and muon trigger thresholds were both 24 GeV throughout the year. The triggers also included loose isolation criteria.

5.3.2 Object Selection

The definitions of electrons, muons and taus used in the analysis is tightened before making the event preselection. Above the threshold of a trigger, there is a gradual increase in efficiency before plateauing. The lepton p_T requirements are tightened, selecting leptons in the plateau region and above only. This is done to simplify any corrections needed due to trigger inefficiencies. In 2011 data, the minimum p_T requirements for electrons and muons are 25 and 22 GeV respectively. In 2012, both electrons and muons have an offline p_T threshold of 26 GeV.

The identification criteria are also tightened. The segment tagged muons are dropped, leaving only combined muons. Furthermore, electrons are required to pass the `tight` working point (70–75% signal efficiency) identification criteria [42]. Finally, the electron veto for taus is tightened to the `medium` working point (85% signal efficiency) for $e\tau_{\text{had}}$ events [49].

It is also required that the leptons are isolated in the detector. Leptons with activity in a surrounding cone are generally indicative of semi-leptonic hadron decays. Isolated leptons, however, are a characteristic of more interesting electroweak decays. The isolation variables are calculated by dividing the energy deposited in a cone with radius ΔR around the objects by the objects total energy. This process can be done for both calorimeter energy deposits and ID track based p_T measurements.

In this analysis, electrons and muons have identical isolation requirements. For the track based isolation, a cone with $\Delta R = 0.4$ is used. A smaller cone is used for the calorimeter isolation, with $\Delta R = 0.2$. In both cases, the total energy in the

surrounding cone must be no more than 6% of the energy of the associated light lepton.

5.3.3 Event Preselection

Before categorising events, a preselection is made which selects good quality events with the correct experimental signature. Any events where the ATLAS detector had significant hardware issues must be ignored. Periods of data taking are checked against the good runs list (GRL), produced centrally by the ATLAS data quality group. To veto cosmic rays and other such events, a selection is made requiring events to have at least one primary vertex. This primary vertex is required to have four or more associated tracks.

Events are also required to have exactly one light lepton and one hadronic tau identified in the event, satisfying the requirements specified above. Furthermore, the two particles should have opposite charges as the Higgs is a neutral boson.

The MMC must have found a valid solution for an event to be considered in the analysis. The MMC mass is an important discriminant and background events have a lower MMC efficiency than signal ones. A cut is made on the transverse mass of the light lepton and the E_T^{miss}, m_T, as this allows the contribution from W+jets events to be reduced. The m_T is defined by the formula

$$m_T = \sqrt{2p_T^{\ell} E_T^{miss}(1 - \cos\Delta\phi)}, \tag{5.3}$$

where $\Delta\phi$ is the gap in ϕ between the lepton and E_T^{miss}. The cut placed on the transverse mass is $m_T < 70$ GeV. Finally, events containing a b-tagged jet with $p_T > 30$ GeV are vetoed to reduce the contamination from top background processes.

Following the application of these loose selection cuts, the events are ready for categorisation. The preselection is summarised in Table 5.2.

5.3.4 Event Categorisation

The analysis categories target two different event topologies. The first category is designed for VBF events. These types of events are characterised by two high energy jets widely separated in η. The second category is named the boosted category. It selects events where the Higgs boson is heavily boosted, typically recoiling from a single high-p_T jet in the opposite direction. In such events, the Higgs often has a large transverse momentum.

The expected event yields for the signal and background processes in each category are summarised in Table 5.3.

Preselection
Event not flagged as bad by GRL
At least one good primary vertex
Pass low-p_T single lepton trigger
Exactly one isolated lepton and one `medium` τ_{had} with opposite signs
$p_T^e > 25$ GeV or $p_T^\mu > 22$ GeV (2011), $p_T^\ell > 26$ GeV (2012)
$p_T^\tau > 20$ GeV
Valid MMC solution
$m_T < 70$ GeV
No b-tagged jets with $p_T > 30$ GeV

Table 5.2: Summary of the preselection cuts used in the analysis [52].

VBF Category

VBF events have two high energy jets widely separated in η. The analysis requires a leading jet with $p_T > 50$ GeV and a subleading jet with $p_T > 30$. The two jets must have $\Delta\eta > 3.0$. Events which fail this jet selection are considered by the boosted category selection.

VBF events are also required to have a visible mass, m_{vis}, larger than 40 GeV. The visible mass is the invariant mass of the tau and light lepton system in the event. Events which pass the jet requirements but fail the cut on m_{vis} are not considered in the boosted category.

Boosted Category

Any events which fail the VBF category's jet requirements are considered for the boosted category. The only requirement made by the boosted category is $p_T^H > 100$ GeV.

5.4 Control Regions

Control regions are areas of the phase space which are selected to be pure in a specific background process. They must also have low numbers of expected signal events to be of use. Control regions are used in this analysis for two primary purposes.

The first is for calculating the fake factors used in the modelling of jet to tau fakes in the analysis. These are discussed in Section 6.4 but are summarised here in Table 5.4.

The second collection of control regions are used to validate the background model. Each of these control regions are rich in a different background process.

Process	VBF	Boosted
Fake background	1680 ± 50	5640 ± 160
$Z \to \tau\tau$	877 ± 29	6210 ± 170
Top	82 ± 15	380 ± 50
$Z \to \ell\ell(\ell \to \tau)$	54 ± 26	200 ± 50
Diboson	63 ± 11	430 ± 40
ggF $H \to \tau\tau$	16 ± 6	60 ± 20
VBF $H \to \tau\tau$	31 ± 8	16 ± 4
WH $H \to \tau\tau$	0.6 ± 0.4	9.1 ± 2.3
ZH $H \to \tau\tau$	0.16 ± 0.07	4.6 ± 1.2
Total background	2760 ± 40	12860 ± 110
Total signal	48 ± 12	89 ± 26

Table 5.3: Expected event yields for each signal and background process in the two analysis categories, normalised to the $\sqrt{s} = 8$ TeV dataset [52]. The given uncertainty on each component is systematic only, while the combined statistical and systematic uncertainty is quoted for the total background and total signal.

Control Region	Definition
QCD Multijet	Signal selection, lepton must fail signal track isolation requirements
W+jets	Signal selection, $m_T > 70$ GeV
Z+jets	Signal selection, two OS same flavour leptons, $80 < m_{\ell\ell} < 100$ GeV
Top	Signal selection, require one or more b-tagged jets

Table 5.4: Control region definitions used for calculating the fake factors utilised in the modelling procedure of the jet to tau fakes background [52].

Control Region	Definition
Fakes	Signal selection, same sign events
W+jets	Signal selection, $m_T > 70$ GeV
Z+jets	Signal selection, two OS same flavour leptons, $61 < m_{\ell\ell} < 121$ GeV
Top	Signal selection, $m_T > 70$ GeV, require one or more b-tagged jets
$Z \to \tau\tau$	Signal selection, $m_T < 40$ GeV, $m_{\tau\tau}^{\mathrm{MMC}} < 110$ GeV

Table 5.5: Control region definitions used to validate the modelling [52]. It should be noted that the $Z \to \tau\tau$ CR is a subset of the signal region. All other CRs are orthogonal to the SR.

Producing and inspecting plots in these regions allows the validity of the background model to be verified for each process individually. These regions are summarised in Table 5.5.

Additionally, a number of other control regions are defined for calculating correction factors. An example of such a region is the region used to correct the dijet distributions for the $Z \to \ell\ell$ MC samples. Such regions are described in the text wherever relevant.

Chapter 6

Signal and Background Modelling

6.1 Introduction

The topic of this thesis is a search for the Higgs decaying to a pair of tau leptons, where one tau decays hadronically and the other decays to a light lepton. The tau decays include the production of neutrinos, which lead to missing energy deposits in the ATLAS detector. Furthermore, jets are used to help identify signal events (discussed in detail elsewhere). This means a wide variety of particles must be considered in the analysis, with a typical event containing a light lepton, a hadronic jet produced by the tau decay, multiple neutrinos and potentially additional hadronic jets.

This combination of particles mean a variety of background processes must be considered, alongside the $H \to \tau\tau$ signal decays. The methods used to model each process will be discussed in detail in this chapter, starting with the main signal processes before moving onto each of the background groups. Accurate modelling forms an important precursor to being able to identify any excesses observed in the data.

6.2 Higgs Signal Processes

Three Higgs production processes are considered in the analysis, introduced earlier in Section 1.9. These are gluon-gluon fusion (ggF), vector boson fusion (VBF) and vector boson associated production (VH), with either a W or Z boson. Cross sections and Feynman diagrams for each process can be seen in Figures 1.2 and

Signal Process	MC Generator	$\sigma \times \mathcal{B}$ (pb) $\sqrt{s} = 8$ TeV	Perturbative Order of Cross Section Calculation
ggF	POWHEG [53, 54, 55, 56] + PYTHIA8 [57]	1.22	NNLO+NNLL
VBF	POWHEG + PYTHIA8	0.100	(N)NLO
VH	PYTHIA8	0.0455 (WH) 0.0262 (ZH)	NNLO

Table 6.1: Details of the signal processes [52]. Showing the MC generator used to produce each signal sample, production cross section multiplied by the $H \to \tau\tau$ branching ratio for $\sqrt{s} = 8$ TeV and $m_H = 125$ GeV, and the perturbative order of the QCD calculation.

1.3 respectively. Furthermore, each process is simulated for a range of Higgs mass values at 5 GeV intervals. The range is 100–150 GeV, with each process and mass point combination modelled by a different Monte Carlo sample.

Relevant details of each of the signal processes are summarised in Table 6.1. Although non-zero, the contribution from $t\bar{t}H$ production is determined to be negligible [52].

The production processes are simulated using POWHEG [53, 54, 55, 56] coupled to PYTHIA8 [57] for both ggF and VBF. For each process, POWHEG's parton distribution functions (PDF) use the NLO CT10 [58] parameterisation.

The ggF sample is normalised to the next-to-next-to-leading-order (NNLO) cross section from the perturbative QCD calculation [59, 60, 61, 62, 63, 64]. Soft-gluon resummation is accounted for up to next-to-next-to-leading logarithm (NNLL) terms [65]. Next-to-leading-order (NLO) electroweak corrections are also applied [66, 67].

In addition to this overall normalisation, a Higgs p_T dependent reweighting is applied to the ggF sample. The reweighting corrects the inclusive Higgs p_T spectrum to match the NNLO QCD calculation using the HRES2.1 software [68], including NNLL resummation corrections. HRES2.1's calculations are used as they account for finite top and bottom quark masses [68, 69] and include dynamical renormalisation and factorisation scales, $\mu_R, \mu_F = \sqrt{m_H^2 + p_T^2}$.

An additional reweighting is applied to ggF events with two or more jets. The Higgs p_T spectrum for these events is reweighted to match that of MINLO HJJ calculations [70].

The overall reweighting is derived so that the inclusive Higgs p_T spectrum agrees with HRES2.1, the p_T spectrum for events with two or more jets matches that of MINLO HJJ and the jet multiplicities match (N)NLO predictions from

JetVHeto [71, 72, 73].

The overall VBF sample normalisation is taken from the NLO perturbative QCD calculation. An approximate NNLO correction is applied [74], as is an NLO electroweak correction [75, 76, 77].

The VBF sample also undergoes a Higgs p_T dependent reweighting process. In this case, the reweighting is due to the p_T dependence of the NLO electroweak corrections to the production cross section. The spectrum is reweighted to that of the Hawk calculation, as this calculation takes into account the p_T dependence of the correction [75, 76]. While the difference is of the order of 1% for low Higgs p_T values, at $p_T = 300$ GeV the correction is approximately 20% [78].

Samples for the final Higgs production process considered, VH, are generated using the Pythia8 with the Cteq6L1 PDF parameterisation [79]. These are normalised to cross section calculations which include NNLO perturbative QCD terms [80] and NLO electroweak corrections [81]. No p_T dependent reweighting is required for these samples.

The taus in all signal samples are decayed and hadronised using Pythia8.

Following the normalisations and reweightings described above, the final steps in the MC simulation process are the addition of pile-up events and detector simulation. Events with no hard interaction are simulated to represent pile-up. These are known as minimum-bias interactions and are simulated using Pythia8 with the AU2 tune [82]. These events are overlaid on the signal events, with pile-up being considered for the same and neighbouring bunch crossings. Following this, the combined event is passed through a process known as full simulation [83]. Here, the Geant4 program [84] is used to simulate the response of the ATLAS detector to the event. Finally, the fully simulated sample is passed through the same reconstruction process as data.

6.3 $Z/\gamma^* \to \tau\tau$

The $Z/\gamma^* \to \tau\tau$ process forms an irreducible background to the search for $H \to \tau\tau$ (for the remainder of this document, all Z/γ^* decays will be written as Z for simplicity). Irreducible means that the observed decay products are the same for both the signal and background processes. This feature, coupled with the large production cross section for Z bosons at the LHC [85], mean this background must be accurately modelled for the analysis to be at all viable.

The modelling of this background in the ATLAS experiment is complicated by two factors. The first is that ATLAS Monte Carlo simulation is found to imper-

fectly model Z+jets processes. If used, the samples must be carefully reweighted to agree with data observations. One alternative is to use a $Z \to \tau\tau$ pure sample, taken from data, to model the background in the signal regions. However, due to the irreducible nature of the background, it is not possible to find a sufficiently signal depleted region for this purpose. For these reasons, a somewhat novel technique is required to model this background in the analysis.

This method takes $Z \to \mu\mu$ events from data as input. The muons are removed from the events and simulated taus are embedded in their place. For this reason, the technique is referred to as embedding. A more detailed description can be found in the following paragraphs.

The first stage in the embedding procedure is finding a highly pure $Z \to \mu\mu$ sample in data. The clean dimuon event signature means efficiently selecting $Z \to \mu\mu$ events is reasonably straightforward.

At trigger level, events passing either the isolated single muon trigger or the dimuon trigger are selected. Here, isolated refers to each muon's tracks and energy deposits in the detector. That is to say, there must be low amounts of activity in a cone surrounding each of the muons in the ID and calorimeters. The trigger coverage is extended to lower p_T by the use of the dimuon trigger as the trigger p_T threshold is lower for both muons.

The offline selection requires events with two isolated, oppositely charged, reconstructed muons. The exact isolation requirement is the sum of the p_T of all tracks within a ΔR of 0.2 of the muon track must be less than 20% of the muon track's p_T. A cut on the dimuon invariant mass is also made, requiring $m_{\mu\mu} > 40$ GeV.

Following the selection of the events, the next step is the removal of the muons and their associated detector deposits. Removing the ID and MS tracks is trivial, any reconstructed tracks matched to the identified muons are simply removed. Subtracting the calorimeter deposits, however, is a more technical procedure.

First, a $Z \to \mu\mu$ event must be generated. The kinematics of the two generated muons are forced to be identical to those of the reconstructed muons in the data event. This generated event is used to estimate the energy deposited in the calorimeter by the muons, and the result is subtracted from the calorimeter towers in the data event. The fact that muons are minimally ionising particles means this procedure is sufficiently accurate.

The third step is the addition of a $Z \to \tau\tau$ decay to the event. The decay is simulated using TAUOLA [86]. Again, the taus are generated with kinematics identical to those of the original muons. However, a correction is applied to account

for the difference in mass between muons and taus. Care is also taken to ensure the polarisations and spin correlations of the two taus are accurate [87]. The ID tracks and calorimeter deposits from the simulated decay are then added to the data event.

The fact that the selected events depend on the muon trigger has the potential to bias the embedding sample. As a final step, each embedded event is weighted to correct for inefficiencies in the muon trigger.

While the process may be slightly complex, the modelling improvements provided make it worthwhile. Using this method means the properties of the Z boson, jets, underlying event and pile-up are all taken directly from data. It is, however, very important to thoroughly validate the various steps of the embedding process. The validation is performed in three steps, each probing a different aspect of the procedure.

The first stage of validation aims to probe the accuracy of the muon energy subtraction in the calorimeters. An additional embedding sample is produced, embedded with muons rather than taus. A comparison is made between the isolation cones around the muons, before and after the embedding. This probes the effect of the subtraction of the original muons and the addition of the simulated ones on the calorimeter energy deposits.

The comparison can be seen in Figure 6.1. The x-axis represents the total transverse energy deposited in the calorimeter, in a cone of radius $\Delta R = 0.3$ about the muon. Thus it shows the energy originating from both the muon and the underlying event. The plot shows good agreement within uncertainties, demonstrating the muon subtraction and embedding procedure does not bias the embedding sample.

The second method of validation relates to verifying the embedding of the taus in the event. This step can be isolated by comparing $Z \to \tau\tau$ MC with taus embedded into $Z \to \mu\mu$ MC. Various kinematics are compared, an example is shown in Figure 6.2. While the quality of the modelling at higher invariant masses shows some variability, it can be seen that the embedding uncertainties encompass the variation. All distributions agree within the uncertainties, confirming that taus can successfully be embedded into events.

The third and final validation technique aims to help authenticate the entire embedding procedure. As the analysis is dominated by $Z \to \tau\tau$, simply comparing between data and background in analysis plots will help verify the procedure. Plots of this nature can be seen throughout this work with the data typically well modelled by the background model, including the embedding component.

This threefold validation appears to fully justify the use of embedding samples. It should be noted that $Z \to \tau\tau$ MC samples are available, any situations

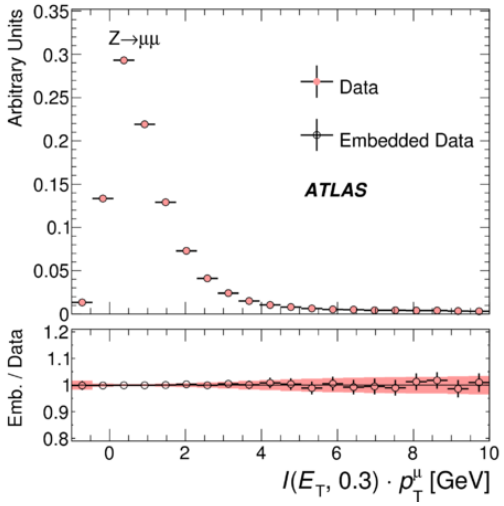

Figure 6.1: Total transverse energy deposited in the calorimeter about muons, before and after muon embedding in data $Z \to \mu\mu$ events [52]. A cone of radius $\Delta R = 0.3$ is used for the comparison. The variable $I(E_T, \Delta R)$ is defined as the total transverse energy deposited in the calorimeter, in a cone of radius ΔR about the muon, divided by the transverse momentum of the muon.

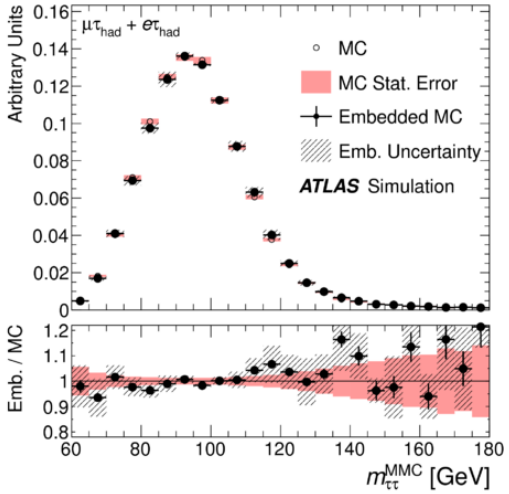

Figure 6.2: The ditau invariant mass, calculated using the Missing Mass Calculator (MMC), compared between $Z \to \tau\tau$ MC and $Z \to \mu\mu$ MC embedded with taus [52]. The MMC is discussed in Section 5.2.7.

where these are used over embedding will state this explicitly. These samples are identical in construction to the $Z \to \ell\ell$ MC samples, described in Section 6.5.1

6.4 Jets Misidentified as Taus

This background category refers to events which pass the selection criteria due to a jet being misidentified as a hadronic tau. A number of processes contribute, namely QCD multijet, W+jets, Z+jets and top production (which is dominated by semileptonic $t\bar{t}$ events). Events of this kind are referred to as jet to tau ($j \to \tau$) fakes. In this analysis, all jet to tau fakes are modelled using the same technique, known as the fake factor (FF) method.

In ATLAS, hadronic jets are passed through a boosted decision tree (BDT) based tau classifier [49]. Higher scoring jets are tagged as tau candidates, with `loose`, `medium` and `tight` identification (ID) levels. A more in depth discussion of tau reconstruction and identification is given in Section 5.2.4.

The FF method estimates the jet to tau fake background using data events where the tau candidate fails to pass the `medium` ID threshold (candidates passing this threshold are used as taus in the analysis). A lower cut is also placed on the tau classifier BDT score, meaning tau candidates must score higher than 0.7 times the `loose` ID threshold to be considered in the FF method. Finally, similar to taus used in the analysis, each tau candidate is required to have one or three associated charged tracks in the inner detector. The tau candidates in this window are referred to as anti-taus from this point forwards.

The `medium` and `loose` ID thresholds vary as a function of the tau p_T. This is to ensure that the ID efficiencies remain roughly constant across the p_T spectrum. The 0.7×`loose` threshold was chosen as a compromise of modelling and statistics. Lowering the threshold allows more events to pass at the cost of poorer modelling (the anti-taus are less tau like). Conversely, raising the threshold improves modelling but the statistical uncertainty increases.

As the numbers of jet to tau fakes above and below the `medium` ID threshold are not identical, factors must be calculated to transfer one to the other. This is from where the fake factor method derives its name.

The FF calculations use data events in four control regions (CRs), one for each of the four backgrounds incorporated into the method. The fake factors are calculated as the ratio between the number of taus (passing the `medium` ID requirement) and anti-taus in each control region. The factors are binned in p_T of the tau candidate and calculated separately for taus with one or three charged tracks.

The control regions used for the calculations are similar to the signal categories introduced in Section 5.3.4, bar some minor changes. The factors are calculated separately for each of the signal categories. This is due to the fake factors exhibiting a dependence on the event topology and the varying background composition in each category. All four control regions utilise data rather than simulated samples.

The gluon dominated QCD multijet CR is defined by inverting the isolation requirements on electrons and muons. The region with high transverse mass, $m_T > 70$ GeV, is selected to create the W+jets CR (see Section 7.3.1 for a full definition of m_T). The Z+jets CR is defined by requiring events to contain two same flavour OS leptons, with $80 < m_{\ell\ell} < 100$ GeV. Finally, for the top CR, the veto on events containing b-tagged jets is inverted. Each of these cuts creates a sample enriched in the relevant background process. These regions are then used to calculate the fake factors.

Before being applied to anti-taus taken from data, the fake factors from each background process must be combined. This is because the process responsible for an event cannot be known on a case by case basis in data. The fake factors for each process are combined, as a function of the expected fraction of events from each source for each signal category. This process can be summarised with the equation

$$FF(p_T, n_{tracks}, \text{cat.}) = \sum_{i=\text{bkg.}} R_i FF_i(p_T, n_{tracks}, \text{cat.}), \tag{6.1}$$

where R_i is the relative fraction of events from each of the four background processes.

For the Z+jets and top processes, the relative fractions (R_Z and R_{top}) are estimated using MC samples. Poor MC modelling of jet to tau fakes in W+jets events means R_W is measured in the W+jets anti-tau data CR. This is then multiplied by a transfer factor which accounts for the difference between the control and signal regions (calculated using simulation). Finally, the fraction of QCD multijet events is simply defined with the equation $R_{QCD} = 1 - R_W - R_Z - R_{top}$.

With the fake factors calculated, the final step is to apply them to the appropriate sample of data events. As stated above, this is the collection of anti-taus in each signal category. However, it should be noted that before applying the fake factors, a component must be subtracted from the anti-taus. Using simulation, the number of events where the anti-tau is a real tau (which fails the `medium` tau ID) or a lepton to tau fake is calculated. This is then subtracted from the anti-tau component before applying the fake factors.

Following this subtraction, each remaining event in the anti-tau sample is

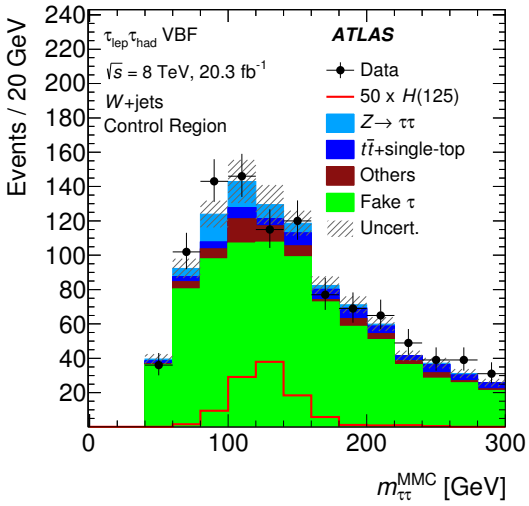

Figure 6.3: The reconstructed ditau invariant mass, $m_{\tau\tau}^{\mathrm{MMC}}$, in the W+jets control region [52]. The large fake tau contribution, modelled using the fake factor method, is seen to be well modelled.

multiplied by the relevant fake factor. The anti-tau samples and FF calculations are kept entirely separate for each signal category. Figure 6.3 shows an example plot from the analysis, with a significant and well modelled contribution from the fake factor method.

One complication of the fake factor method relates to the calculation of the missing transverse energy, E_T^{miss}, in anti-tau events. For reasons relating to the soft term contribution to the E_T^{miss}, the choice of tau can change the value of the E_T^{miss}. A correction was devised which ensures the anti-taus are correctly treated as taus before recalculating the E_T^{miss} for each event. This fix was not possible for 2011 data for technical reasons, therefore an appropriate reweighting was applied to these events to reduce the mismodelling.

6.5 Monte Carlo Background Simulation

The remaining background contributions are modelled using various Monte Carlo generators. One thing each sample has in common, however, is the simulation of pile-up and the detector response. These are identical to those of the signal MC samples, described at the end of Section 6.2. The production cross sections of the background processes (including those modelled by the embedding and fake factor methods) can be seen in Table 6.2.

Background Process	$\sigma \times \mathcal{B}$ (pb) $\sqrt{s} = 8$ TeV
$W(\to \ell\nu), (\ell = e, \mu, \tau)$	36,800
$Z/\gamma^* \to \ell\ell$, 60 GeV $< m_{\ell\ell} <$ 2 TeV	3,910
$Z/\gamma^* \to \ell\ell$, 10 GeV $< m_{\ell\ell} <$ 60 GeV	13,000
VBF $Z/\gamma^* \to \ell\ell$	1.1
$t\bar{t}$	253
Single top: Wt	22
Single top: s-channel	5.6
Single top: t-channel	87.8
$q\bar{q} \to WW$	54
$gg \to WW$	1.4
WZ, ZZ	30
$H \to WW$	4.7

Table 6.2: Cross sections of the background processes at $\sqrt{s} = 8$ TeV [52]. These values are used to normalise the backgrounds modelled using MC simulation. The W and Z/γ^* processes are quoted with the branching ratio to leptons included. The remaining cross sections listed are inclusive.

6.5.1 $Z/\gamma^* \to \ell\ell$

While $Z \to \ell\ell$ events where the tau is faked by an additional jet are modelled using the fake factor method, events where the tau is faked by a lepton are modelled using MC simulation. The nominal MC sample is produced using Alpgen [88] coupled to Pythia8.

Additional samples are produced, filtered to select only VBF events. The VBF filter is applied at event generation level, selecting events with a large dijet mass ($m_{j1,j2} > 200$ GeV) and separation in $\Delta\eta$ between the two leading jets ($\Delta\eta_{j1,j2} > 2.0$). This statistically enriches the VBF phase space in the final MC sample, helping to ensure the accuracy of the modelling. This is necessary due to the targeting of VBF signal events in the analysis.

The VBF filter samples are generated with Alpgen coupled to Herwig [89]. Both Alpgen and Herwig use the Cteq6L1 PDF parameterisation. The nominal and VBF samples are combined into one sample, named the hybrid sample. This is done by discarding events in the nominal sample which would pass the VBF truth filter, before weighting the samples by their respective cross sections and combining them.

Due to an observed mismodelling in dijet variables, a reweighting was derived for the $Z \to \ell\ell$ samples. The correction reweights the $|\Delta\eta_{j1,j2}|$ distribution in MC to match the observations in data. This is done in a $Z \to \ell\ell$ CR, which requires two same flavour OS leptons with $|m_Z - m_{\ell\ell}| < 15$ GeV.

There are also issues with the modelling of the $\ell \to \tau$ misidentification rate. This is corrected for, using additional scale factors to scale the MC samples. The correction is especially important for the $Z \to ee$ samples.

As mentioned above, similar samples are available for the $Z \to \tau\tau$ process. These are produced identically, including the merging of the hybrid sample and application of the dijet reweighting. The lepton to tau fake rate corrections are not necessary however.

6.5.2 Top

Top modelling encompasses four processes. Top quarks can be produced in pairs, this is referred to as $t\bar{t}$ production. A single top can be produced when a $W^{\pm}$ boson decays to one top and one bottom-type quark (s-channel) or when a bottom-type quark absorbs a $W^{\pm}$ (t-channel). Finally, a top quark can be produced in conjunction with a $W^{\pm}$ boson (Wt production).

Aside from jet to tau fakes, modelled using the fake factor method, events with real taus form the major background contribution from top quark events. A small contribution is also expected where the tau is faked by a lepton.

$t\bar{t}$, s-channel and Wt events are modelled by POWHEG+PYTHIA8. The t-channel production sample is produced using AcerMC [90] coupled to PYTHIA6 [91]. AcerMC utilises the CTEQ6L1 PDF parameterisation in ATLAS.

The top backgrounds are normalised using a top enriched control region. This is defined by inverting the m_T cut and b-jet veto used in the signal regions of the analysis. The region is orthogonal to the control region used to calculate top FF.

6.5.3 Diboson

This background contains events where two weak vector bosons are produced. The diboson contribution to the background is entirely modelled using Monte Carlo simulation.

WW events are split into two subsamples, depending on the origin of the $W^{\pm}$ bosons. GG2WW [92] linked to HERWIG is used to simulate the $gg \to WW$ contribution. The CT10 PDF parameterisation is used by GG2WW. Events from the process $q\bar{q} \to WW$ are generated using ALPGEN+HERWIG. WZ and ZZ events are generated using standalone HERWIG.

6.5.4 Others

There are two additional background processes considered in the analysis. The contribution from each is expected to be negligible but they are included for completeness.

The first is $W \to \ell\nu$. This background's contribution is almost entirely from events where a jet fakes the tau, modelled by the fake factor method. ALPGEN+PYTHIA8 is used to generate samples which are used to model the remaining $W \to \ell\nu$ contribution. This is predominantly events where the tau is faked by a light lepton.

Due to the nature of the analysis, $H \to WW$ must be considered a background process. The tiny expected contribution is simulated using techniques identical to how the $H \to \tau\tau$ production mechanisms are modelled, described in Section 6.2.

Chapter 7

Boosted Decision Trees

7.1 Introduction

Boosted decision trees (BDTs) [93] are a type of machine learning multivariate predictor. They can be used to predict whether a specific event is more signal- or background-like. BDTs are used by this analysis as they do not outright discard events, unlike cut-based analyses. This, and the fact that BDTs are able to exploit correlations between variables, allows the sensitivity of the analysis to be improved. Two BDTs are used in this analysis, one optimised for the VBF signal category and the other for the boosted category. The output score from the BDTs, which describes how signal-like each event is, is used as the discriminant in the fit of the analysis.

BDTs are a more advanced form of decision trees. The first stage of developing a decision tree is to split each sample into two subsamples, half for training the decision tree and the other half for testing its performance. Each event in the input sample must be labelled as either a signal or background event. Also, a list of variables must be chosen for use by the decision tree training algorithm. Ideally, each of the chosen variables should provide some separation between signal and background events. However a further benefit of decision trees is that they are also capable of exploiting correlations between the variables.

The training of a decision tree is an iterative process. Each variable is scanned over to find the cut which provides the best separation between signal and background events. The Gini index [94] is the metric used to decide which cut performs best. However, unlike in a cut-based analysis, both subsamples are retained.

This training process is repeated independently on the two subsamples, both the one which passed and failed the previous cut. The process continues, dividing

each subsample in two each time, until every subsample satisfies one or more end conditions. These are: the subsample reaching a high signal or background purity, the subsample containing fewer than a minimum number of events or the decision tree reaching a specified maximum depth.

Once training is complete, one is left with a tree where each fork of the branches is the cut with the best signal-to-background separation. The end nodes (or "leaves") are each labelled as signal or background, depending on their composition. Events of unknown source can then be passed through the cuts of the tree to determine whether they are more likely to be signal or background events.

The training algorithm described above will most likely result in some training events being categorised incorrectly. The use of boosting allows the misclassification rate to be reduced, improving the accuracy of the BDT when applied to new events.

7.2 Boosting

Boosting is a technique where a collection of weak learners, such as decision trees, can be combined to form one strong learner. When boosting decision trees, events misclassified by the first tree are given a higher weight. These reweighted events are then used to train another decision tree. This process is repeated, typically hundreds of times, producing a forest of decision trees.

A weighted combination of decisions from every tree is calculated for each event the BDT is applied to. This gives each event a score, with more signal like events having a higher score. The boosting algorithm used in this analysis is the gradient boosting algorithm [95].

7.3 Training

Before training the BDTs used in this analysis, a number of decisions had to be made. These include selecting the discriminating variables and choosing the BDT parameters (such as minimum leaf size and number of trees in the forest).

A short list of variables is preferred for each of the two BDTs, as adding variables typically leads to additional systematic uncertainties which must be controlled. First, lists of potentially useful variables are drawn up for each of the two categories. Each variable should provide some discrimination against a background process. For example, the VBF category variable long list featured many jet and dijet related variables. These typically provide good discrimination between VBF

signal and non-signal events.

The modelling of each variable must then be studied, with badly modelled variables discarded. This is done to prevent the BDT from targeting artificial features, not present in data events. An example of such a variable is $\Delta\Delta R_{\tau\ell}$. A fit is made of the $\Delta R_{\tau\ell}$ dependence on the Higgs p_T for each category. $\Delta\Delta R_{\tau\ell}$ quantifies the deviation from the expected value of $\Delta R_{\tau\ell}$, from the fit, for an event with a given Higgs p_T. The variable was imperfectly modelled and therefore replaced by $\Delta R_{\tau\ell}$, for a minimal loss in sensitivity.

The next step is to train the BDTs with the preliminary variable lists. This was done using the Toolkit for Multivariate Analysis [96].

This analysis uses a technique named cross-training. The training sample for each category is randomly split in two, with a different BDT being trained using each half. Each BDT is then applied to the other half of the training set, to test its performance. A check is made to ensure that both BDTs have similar performance. When applying the BDTs to data, the data is randomly split into two halves, with one of the cross-trained BDTs applied to each.

When training the BDTs, the MC $Z \to \tau\tau$ samples (including the VBF filtered samples) were used rather than the embedding samples. This is because they provide greatly improved statistics, an important consideration when training a multivariate classifier. For the VBF category, only the VBF Higgs samples are used as signal when training the BDTs. This forces the BDT to more specifically target the VBF event topology. The boosted category uses all four production processes as the signal component during training, with each process weighted to its production cross section. All other processes were modelled as discussed in Chapter 6.

The BDT training parameters are carefully optimised. This was done by calculating the statistics-only sensitivity of the analysis when varying the parameters then selecting the parameters which give the best sensitivity.

Attempts to further trim the variable list are then undertaken. The training procedure provides information on variable correlations as well as a ranking of how effective each variable is at distinguishing between signal and background. Variables are removed one by one, beginning with the highly correlated and lowly ranked variables. If removing a variable only causes a slight worsening of the statistics-only sensitivity, the variable is dropped from future consideration. Variables whose removal significantly degrades the sensitivity are reinstated.

Variable	Description	VBF	Boosted
$m_{\tau\tau}^{\text{MMC}}$	Invariant mass of ditau system, calculated using MMC tool	•	•
m_T	Transverse mass of ℓ and E_T^{miss}	•	•
$\Delta R_{\tau\ell}$	Separation between τ and ℓ	•	•
E_T^{miss} ϕ centrality	See text	•	•
$m_{j1,j2}$	Invariant mass of two leading jets	•	
$\eta_{j1} \times \eta_{j2}$	Product of ηs of two leading jets	•	
$\Delta\eta(j1, j2)$	Difference in ηs of two leading jets	•	
ℓ η centrality	See text	•	
p_T^{total}	Magnitude of vector sum of τ, ℓ, $j1$, $j2$, E_T^{miss} p_T	•	
$\sum \lvert p_T \rvert$	Sum of τ, ℓ, jets p_T		•
p_T^{ℓ}/p_T^{τ}	p_T ratio for ℓ and τ		•

Table 7.1: Discriminating variables used in the BDT for the two analysis categories [52].

7.3.1 Final Variable List

The output of the procedure described above forms the final variable list, used to train the BDTs. The variables of the list will now be discussed in turn, with care taken to describe any unique properties and where their discriminatory power originates from. The final variable list is summarised in Table 7.1.

$m_{\tau\tau}^{\mathbf{MMC}}$

The invariant mass of the ditau system, as reconstructed by the MMC tool. The MMC tool is described in Section 5.2.7. Including the ditau mass may bias the BDTs to prefer resonances with a reconstructed mass around 125 GeV. This means the analysis specifically targets the recently discovered boson, rather than being a more general search.

One alternative would be to remove the mass from the BDTs entirely, then make a two dimensional fit to mass and BDT score. While heavily investigated, this approach was found to be significantly less sensitive than the one dimensional fit to BDT score. The primary reason for this is that the mass information allows the BDTs to more readily distinguish between signal and the irreducible background of $Z \to \tau\tau$.

The variable distribution for the VBF and boosted categories can be seen in the upper-left plot of Figure 7.1 and Figure 7.3 respectively.

m_T

The transverse mass of the light lepton and the E_T^{miss}. This was defined above with the formula

$$m_T = \sqrt{2p_T^{\ell} E_T^{miss}(1 - \cos\Delta\phi)}, \tag{5.3}$$

where $\Delta\phi$ is the gap in ϕ between the lepton and E_T^{miss}.

The variable exhibits particularly good separation between signal events and $W \to \ell\nu$ events where a jet fakes the hadronic tau. In such events, the E_T^{miss} and lepton both originate from the decay of the $W^{\pm}$ boson. This means $\Delta\phi$ is small and hence m_T is large. For signal events the E_T^{miss} is made up of three neutrinos from two tau decays, meaning the angle between the E_T^{miss} and the lepton is typically larger and the m_T is reduced.

The variable distribution for the VBF and boosted categories can be seen in the upper-right plot of Figure 7.1 and Figure 7.3 respectively.

$\Delta R_{\tau\ell}$

The angular separation between the tau and the lepton, defined using the formula

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{7.1}$$

For resonance decays, such as $H \to \tau\tau$ and $Z \to \tau\tau$, $\Delta R_{\tau\ell}$ tends to be small. This is because the tau and lepton originate from the same parent particle. No such restriction applies to the other backgrounds, so these events are more spread across the $\Delta R_{\tau\ell}$ spectrum. This provides reasonable separation between signal events and the fakes background, for both the VBF and boosted categories.

The variable distribution for the VBF and boosted categories can be seen in the middle-left plot of Figure 7.1 and Figure 7.3 respectively.

E_T^{miss} ϕ Centrality

The use of continuous variables is preferred over boolean flags as they give the BDTs improved discriminatory powers.

The E_T^{miss} ϕ centrality has a maximum value of $\sqrt{2}$ when the E_T^{miss} is directly between the lepton and τ in ϕ. It is equal to 1 if the E_T^{miss} has a ϕ coordinate identical to the lepton or τ and is < 1 elsewhere.

The equation for E_T^{miss} ϕ centrality (C_ϕ) is

$$A = \frac{\sin\left(\phi_{E_T^{miss}} - \phi_\tau\right)}{\sin\left(\phi_\ell - \phi_\tau\right)}, \tag{7.2a}$$

$$B = \frac{\sin\left(\phi_\ell - \phi_{E_T^{miss}}\right)}{\sin\left(\phi_\ell - \phi_\tau\right)}, \tag{7.2b}$$

$$C_\phi = \frac{A + B}{\sqrt{A^2 + B^2}}. \tag{7.2c}$$

In $H \to \tau\tau$ (and $Z \to \tau\tau$) decays, the E_T^{miss} will typically be located between the hadronic tau and the lepton. This is because the two tau decays, which are the origin of the hadronic tau and lepton, both contribute to the E_T^{miss}. Therefore the E_T^{miss} ϕ centrality tends towards high values for signal events. This is especially true for events in the boosted category. Therefore this variable gives separation between the signal and the non-resonant backgrounds.

The variable distribution for the VBF and boosted categories can be seen in the middle-right plot of Figure 7.1 and Figure 7.3 respectively.

Dijet Variables ($m_{j1,j2}$, $\eta_{j1} \times \eta_{j2}$ and $\Delta\eta(j1, j2)$)

VBF events are characterised by two high-p_T jets in opposite halves of the detector. In VBF events, two quarks each emit a vector boson and the two bosons fuse to produce the Higgs. The two quarks are from protons travelling in opposite directions in the LHC and are the source of the two high-p_T jets. This origin helps explain the high momenta of the jets as well as their positions in opposite detector hemispheres.

The three dijet variables aim to quantify such features. They tend to take on more extreme values for VBF signal events, providing reasonable discrimination against all backgrounds. The dijet mass ($m_{j1,j2}$) and η gap ($\eta_{j1} \times \eta_{j2}$) are typically large while the η product ($\Delta\eta(j1, j2)$) tends towards large negative values.

The variable distributions for the VBF category can be seen in the lower-left and lower-right plot of Figure 7.1 and the upper-left plot of Figure 7.2.

ℓ η Centrality

The use of continuous variables is preferred over boolean flags as they give the BDTs improved discriminatory powers.

The ℓ η centrality is at a maximum value of 1 when the lepton is directly between the two leading jets in η. This drops to $1/e$ when the lepton is aligned with

one of the two leading jets and decreases further elsewhere. The exact formula is

$$C_\eta = \exp\left(\frac{-1}{\eta_{j1} - (\frac{\eta_{j1}+\eta_{j2}}{2})^2}\left(\eta_\ell - \frac{\eta_{j1}+\eta_{j2}}{2}\right)^2\right). \tag{7.3}$$

VBF signal events typically have higher values, giving some separation against all backgrounds. This is because the Higgs, and hence the lepton, is generally produced centrally in the detector in VBF events. This is coupled with the fact that the two jets are expected to be in opposite halves of the detector. These two factors increase the chance that the lepton will be located between the two jets in η.

The variable distributions for the VBF category can be seen in the upper-right plot of Figure 7.2.

p_T^{total}

The magnitude of the vector sum of the p_T of the tau, lepton and two leading jets plus the E_T^{miss}. The combination of the tau and lepton p_Ts and the E_T^{miss} approximately reconstructs the Higgs candidate. Coupled with the two leading jets, this accounts for all the constituents of VBF production. Hence, the p_T of the Higgs and two jets should balance and the vector sum should be around zero. This gives separation between VBF events and the background processes.

The variable distributions for the VBF category can be seen in the lower plot of Figure 7.2.

$\sum |p_T|$

Sum of the p_Ts of the tau, lepton and jets (if any) of the event. This is effectively a measure of the total activity in the event and is hence expected to be higher for events where the Higgs is boosted by a recoiling jet. The main sensitivity of this variable is due to its correlation with $\Delta R_{\tau\ell}$. More heavily boosted events (and therefore with a higher $\sum |p_T|$) have a lower $\Delta R_{\tau\ell}$, as the boost makes the decay products more collimated. However, the Higgs has a higher mass then the Z^0 boson and this additional energy in Higgs decays decreases the collimating effect. This allows signal events to be more readily distinguished from the $Z \to \tau\tau$ background. The other backgrounds typically have fewer high-p_T objects, providing some additional separation.

The variable distributions for the boosted category can be seen in the lower-left plot of Figure 7.3.

p_T^{ℓ}/p_T^{τ}

The p_T ratio of the lepton and tau. In a $H \to \tau_\ell \tau_{had}$ event, the hadronically decaying tau's signature is a hadronic jet and a neutrino. The leptonically decaying tau produces a lepton alongside two neutrinos. The additional neutrino in the leptonic case suggests that this p_T ratio should typically be slightly below one. This feature provides separation between signal events and the reducible backgrounds.

The variable distributions for the boosted category can be seen in the lower-right plot of Figure 7.3.

7.4 Validation

Due to the complex nature of a BDT based analysis, extensive validation is required. For example, care must be taken to ensure the BDTs have not been overtrained. Overtraining is a phenomenon where the BDT training procedure "learns too much" from the input training sample. BDTs suffering from overtraining perform well on the training sample but struggle when applied to new data.

During the analysis, the modelling of all the BDT input variables was checked in each SR and each of the validation CRs listed in Table 5.5. Additionally, the modelling of the BDT output score was verified in each CR. Plots showing the distributions of the input variables for the VBF and boosted BDTs are shown in Figures 7.1, 7.2 and 7.3. Some of the control region plots can be seen in Figure 7.4.

Each plot is produced using the 20.3 fb^{-1} of data recorded at a centre of mass energy of 8 TeV. Each process is modelled as described in Chapter 6 and all MC samples are normalised to their expected number of events. The combined statistical and systematic uncertainty on the background model is shown as a grey striped band. The data's statistical uncertainty is shown as an error bar on the black data points. For improved visibility, the signal MC is scaled up by a factor of 50 for the SR plots and 10 for the CR ones.

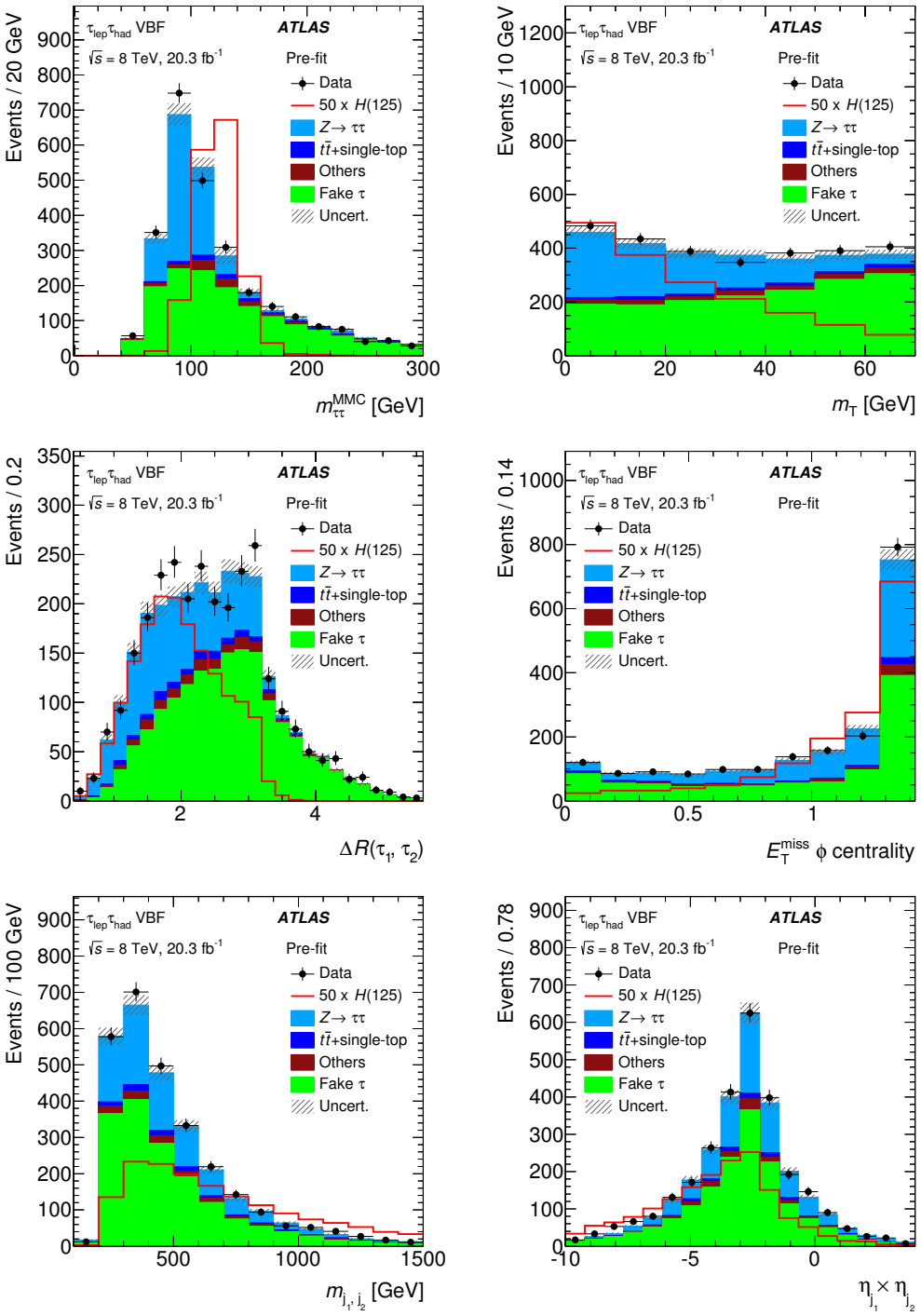

Figure 7.1: Input variables used by the VBF category BDT. From left to right, top to bottom: $m_{\tau\tau}^{\mathrm{MMC}}$, m_T, $\Delta R_{\tau\ell}$, E_T^{miss} ϕ centrality, $m_{j1,j2}$ and $\eta_{j1} \times \eta_{j2}$.

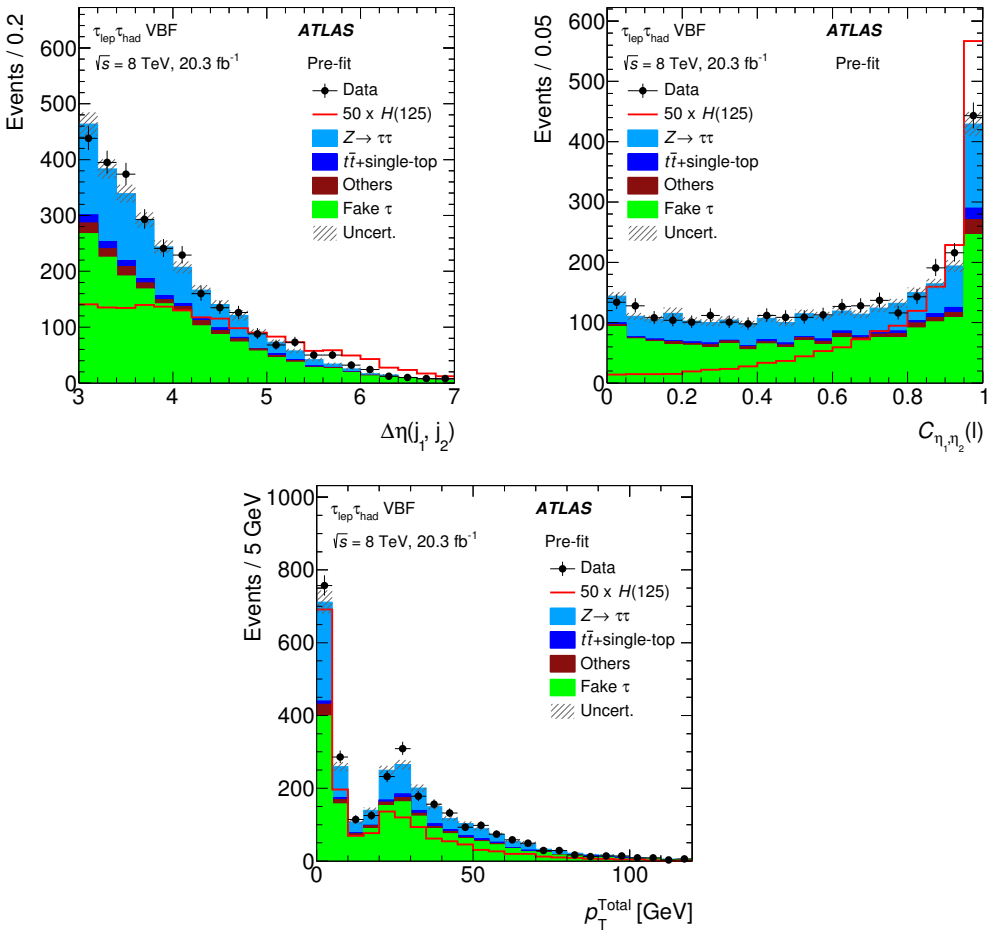

Figure 7.2: Input variables used by the VBF category BDT. From left to right, top to bottom: $\Delta\eta(j1, j2)$, ℓ η centrality and p_T^{total}.

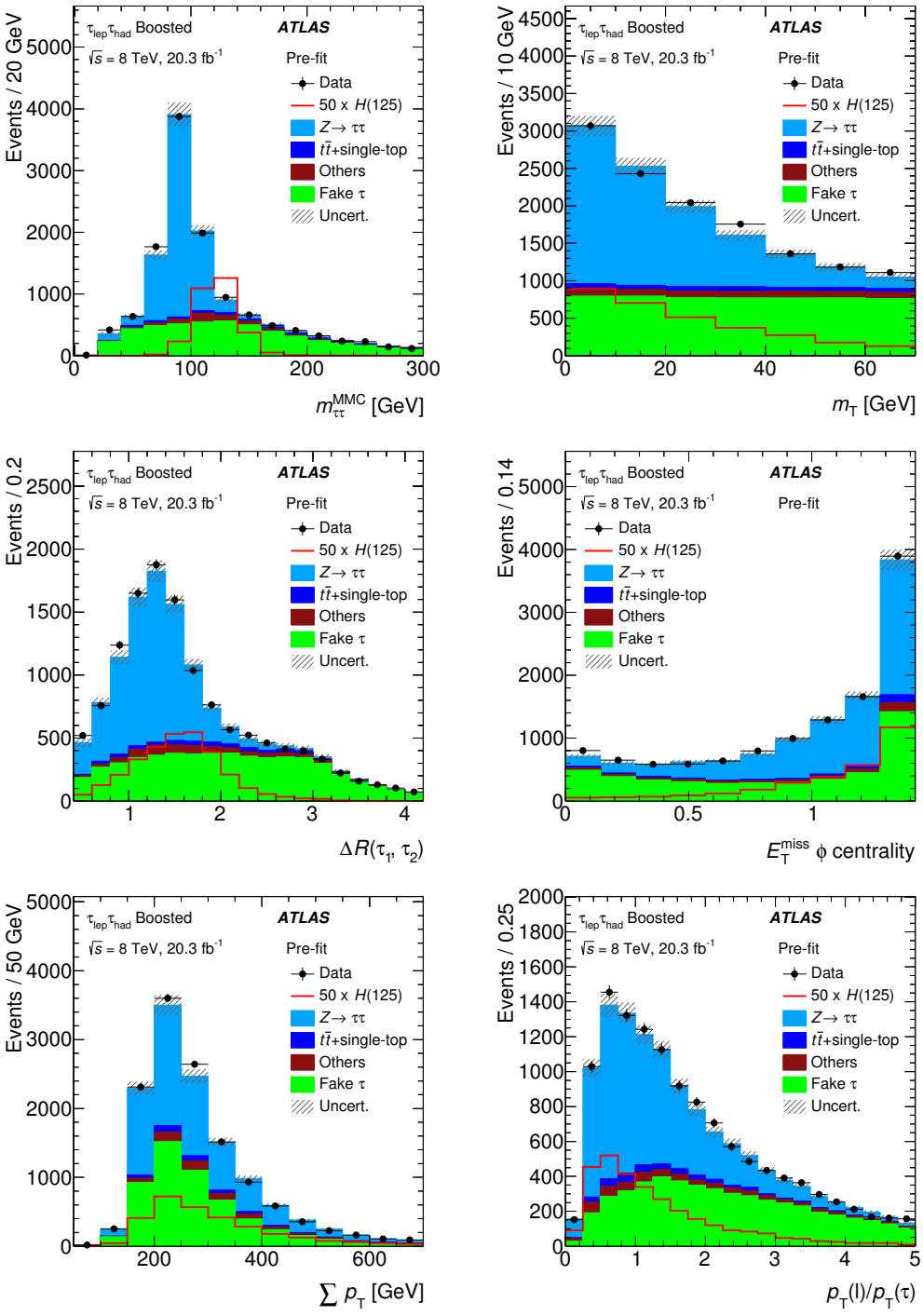

Figure 7.3: Input variables used by the boosted category BDT. From left to right, top to bottom: $m_{\tau\tau}^{\mathrm{MMC}}$, m_T, $\Delta R_{\tau\ell}$, E_T^{miss} ϕ centrality, $\sum |p_T|$ and p_T^{ℓ}/p_T^{τ}.

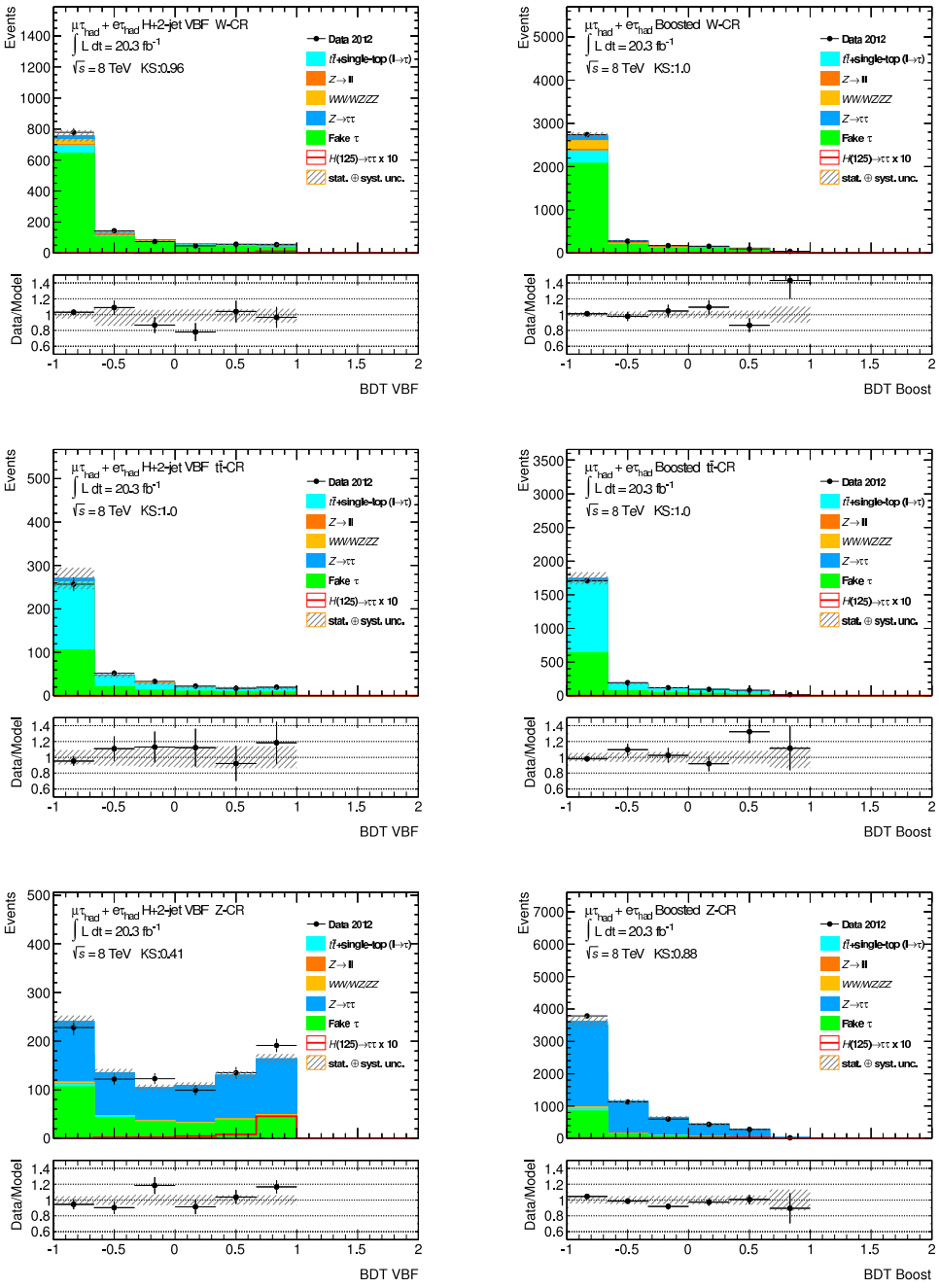

Figure 7.4: BDT score distributions in three of the analysis control regions. VBF CRs are on the left with boosted on the right. From top to bottom, the plots show the W+jets, top and $Z \to \tau\tau$ CRs, as defined in Table 5.5.

Chapter 8

Fitting

8.1 Introduction

The BDT output in the two signal categories is the final discriminant used in the analysis. It is used as the input to a maximum-likelihood fit, simultaneously fitting in all categories and channels. All fitting is performed using the `HistFactory` software [97].

The expected number of events in each bin of the BDT score distributions is predicted by the signal and background models, described in Chapter 6. The models have a number of parameters which can be varied in the global fit, so that the model describes the observed data as accurately as possible [98]. These parameters are divided into two categories—parameters of interest (PoIs) and nuisance parameters (NPs)—depending on their origin.

The name parameters of interest accurately describes their role in the fit. The general aim of the fit model is to accurately measure these parameters and their uncertainty. This analysis has one PoI, the signal strength parameter which is represented by the symbol μ. The signal strength parameter is given in units of the Standard Model signal expectation. For example, a best-fit value of $\mu = 0$ represents no evidence for the Higgs boson while $\mu = 1$ would signify a Higgs signal with the expected Standard Model cross section. The uncertainty on μ is also important to quantify the significance of any measurement. The SM cross sections used are given in Table 6.1.

The nuisance parameters, on the other hand, represent parameters whose values are of limited physical interest. Each NP quantifies a systematic uncertainty, allowing the data to constrain the uncertainties. The NPs are discussed in more detail in Section 8.2 below. The vector $\boldsymbol{\theta}$ is used to represent the NPs.

The binned likelihood function is constructed by multiplying the Poisson probability terms of every bin in each histogram used in the fit

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{i \in \text{bins}} \frac{\lambda_i^{n_i}}{n_i!} e^{-\lambda_i}. \tag{8.1}$$

Here, λ_i is the expected number of events in bin i and n_i is the observed number of events. This incorporates all bins in the four signal region BDT histograms (two categories $\times$ two data periods) plus four single bin histograms which quantify the total number of events in each of the top control regions (again, two categories $\times$ two data periods). The latter are included to help constrain the overall normalisation of the top related backgrounds, a free parameter in the global fit.

The binned likelihood function can be used to construct the test statistic

$$q_\mu = -2 \ln \frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})}. \tag{8.2}$$

Here, $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the parameters which maximise the likelihood and $\hat{\hat{\boldsymbol{\theta}}}$ maximises the likelihood for a specified value of μ. The larger the value of q_μ, the greater the incompatibility of the specified μ with the observed data.

The level of disagreement can be better quantified by calculating the p-value,

$$p_\mu = \int_{q_{\mu,\text{obs}}}^{\infty} f(q_\mu|\mu) dq_\mu. \tag{8.3}$$

The p-value is the probability of obtaining a measurement equal or more extreme than the one observed, given the signal strength hypothesis μ. $f(q_\mu|\mu)$ is the probability density function (PDF) of q_μ for the signal strength μ. By setting $\mu = 0$, the compatibility with the background-only hypothesis can be calculated (and excluded for suitably small values of p).

The shape of $f(q_\mu|\mu)$ cannot be known exactly but it can be approximated in certain conditions. In the asymptotic limit of a large number of events, the PDF can be approximated by an analytic function. This function is independent of the NPs, depending only on the variance of the best-fit value of the signal strength parameter, $\hat{\mu}$ [98].

Finally, this p-value can be converted into a significance, Z. Z is the number of standard deviations above the mean which has an upper-tail probability equal to the p-value, for a normal distribution. This can be calculated using

$$Z = \Phi^{-1}(1 - p). \tag{8.4}$$

Here, Φ^{-1} is the inverse cumulative distribution function for the normal distribution. In particle physics, a $Z \geq 5$ rejection of the background hypothesis is required in order to claim evidence for a discovery [98].

8.2 Nuisance Parameters

Each of the analysis NPs describes a different uncertainty in the analysis. The uncertainties fall into three major subcategories: theoretical, experimental and statistical uncertainties. They also come in two varieties, normalisation and shape uncertainties.

A normalisation uncertainty is the simpler of the two varieties, with the total normalisation of any affected processes able to vary up or down. An example of such a variation is the uncertainty on the integrated luminosity recorded by the ATLAS detector. All MC samples are scaled to the integrated luminosity of the dataset used in the analysis. This is the full Run 1 ATLAS dataset, 4.5 fb^{-1} at $\sqrt{s} = 7$ TeV and 20.3 fb^{-1} at $\sqrt{s} = 8$ TeV. This uncertainty therefore applies to all MC samples used in the analysis, allowing their overall normalisations to vary by $\pm 2.8\%$ ($\pm 1.8\%$) for the 2012 (2011) dataset.

Shape uncertainties, on the other hand, allow the contribution from individual events to change. This can be due to the event weights varying as a function of the event variables. For example, the MC weights which correct for any MC trigger mismodellings are a function of η and ϕ. Additionally, entire events may enter or leave the event selection due to systematic variations in their event variables. The jet energy scale (JES) uncertainties, for example, allow the energies of each jet in an event to vary. This can lead to events no longer passing the selection cuts for the VBF category. Therefore, shape uncertainties cause bin-by-bin variations of the expected number of events in the BDT score histograms.

Before being passed as input to the fit, each nuisance parameter undergoes a pruning and smoothing procedure. The process and the reasons behind it are discussed in Section 8.2.1. Following that, the sources of all NPs of the analysis will be described.

The bulk of the uncertainties can be assigned to three major categories. These are the experimental systematic uncertainties (Section 8.2.2), the background modelling uncertainties (Section 8.2.3) and the theoretical uncertainties (Section 8.2.4). Each is discussed in turn.

The discussion below centres on the uncertainty calculations for the signal processes. Where appropriate, equivalent background uncertainties are calculated

using the same techniques. The impact of the uncertainties are summarised in Table 8.1.

8.2.1 Pruning and Smoothing

Following the application of the final selection cuts, the number of events remaining for each process can be fairly low in statistics. This is especially true for the high BDT score bins where the background processes should be heavily suppressed. This carries the risk that any shape systematic uncertainties are dominated by statistical noise rather than systematic variations. The pruning and smoothing processes outlined here form a conscious effort to systematically prune away or smooth out such noisy variations.

Pruning (Pre-smoothing)

This initial pruning procedure is applied only to shape systematics where the event variables are able to vary. It is not applied to systematics that only allow the individual event weights to change. For each of these shape systematics, a χ^2 test is made between the upwards and downwards fluctuation versus the nominal case. This is done separately for every process which the shape systematic applies to. When calculating the χ^2, only the larger of the nominal and varied statistical uncertainty is considered as the two cases are statistically correlated.

The probability of observing such an extreme χ^2 is calculated and the shape variation is retained if its p-value is below 0.98. If not, the shape variation of the systematic is ignored but the overall effect on the process normalisation is retained as a normalisation systematic.

Smoothing

Following the first round of pruning, all shape systematics undergo the smoothing procedure. Firstly, the ratios between the upward and downward variations and the nominal case are calculated for each bin. The ratios are used as the BDT score histograms are expected to exhibit some shape variation, whereas the ratios should be smoothly varying. Then, these histograms are smoothed using ROOT's `TH1::Smooth(1)` method [99]. The upwards and downwards variations are finally recovered by multiplying the nominal histogram by the two smoothed ratio histograms.

Pruning (Post-smoothing)

Finally, another round of pruning is performed. For the shape uncertainties, the significance of the variation in each bin is calculated separately for each process using the formula

$$S_i = \frac{|u_i - d_i|}{\sigma_i^{\text{tot}}}. \tag{8.5}$$

u_i and d_i are the upward and downward systematic variation in bin i. σ_i^{tot} is the total statistical uncertainty for the background estimation in the bin. The pruning algorithm requires $S_i > 0.1$ for at least one bin for each process, otherwise the shape uncertainty is neglected for that process.

Lastly, for each process, any normalisation uncertainty which causes less than a 0.5% variation is ignored. This speeds up the fitting process and is deemed acceptable as the variation from such uncertainties can easily be "absorbed" by other, larger, uncertainties. Normalisation uncertainties smaller than 0.5% appear in Table 8.1 as the uncertainties are quoted for the combination of all processes. An uncertainty greater than 0.5% on one process may be combined with other processes where the uncertainty does not apply, leaving a final value below 0.5%.

8.2.2 Experimental Systematic Uncertainties

The experimental uncertainties primarily represent unknowns of the ATLAS detector. Effort is made to accurately model the detector properties in MC simulation but this process introduces a number of systematic uncertainties. Data driven background modelling techniques allow the reduction of such uncertainties by taking the event variables directly from data events. However, the complexities of the methods do mean that additional systematic uncertainties are required. These are described in Section 8.2.3.

The $Z \to \tau\tau$ background, modelled using the embedding technique, takes all but the tau decays from data. This means that only the tau related experimental systematic uncertainties are considered for the embedding sample. This removes the need for uncertainties on the jets, leptons and E_T^{miss}, greatly simplifying the systematic treatment.

The fake factor method, used to model background events where the tau is faked by a jet, also simplifies the systematic uncertainty treatment. As all components of the event come directly from data, no experimental uncertainties are needed. The method does introduce its own uncertainties however, these are described in Section 8.2.3.

	VBF Signal	VBF Background	Boosted Signal	Boosted Background
Experimental				
Luminosity	± 2.8	± 0.1	± 2.8	± 0.1
TES †	± 2.4	± 1.3	± 2.4	± 0.9
JES and JER †	$^{+9.5}_{-8.7}$	± 1.0	± 3.9	± 0.4
E_T^{miss} soft scale and res.	$^{+0.8}_{-0.3}$	± 0.2	± 0.4	< 0.1
Lepton id. and trigger	± 1.8	± 0.5	± 1.8	± 0.8
Tau id.	± 3.3	± 1.2	± 3.3	± 1.8
b-tagging	< 0.1	± 0.2	± 0.4	± 0.2
Background Model				
Embedding †	—	± 1.5	—	± 1.2
Fakes modelling †	—	± 2.6	—	± 2.6
Theoretical				
Higher order QCD corr. †	$^{+9.7}_{-7.6}$	± 0.2	$^{+19.3}_{-14.7}$	± 0.2
UE/PS	± 3.8	< 0.1	± 2.9	< 0.1
Generator modelling	± 2.7	< 0.1	± 1.3	< 0.1
EW corr.	± 1.3	< 0.1	± 0.4	< 0.1
PDF	$^{+3.9}_{-3.6}$	± 0.2	$^{+6.6}_{-6.1}$	± 0.2
BR ($H \to \tau\tau$)	± 5.7	—	± 5.7	—

Table 8.1: The impact of systematic uncertainties for signal and background processes in the various signal regions, expressed as a percentage variation [52]. The quoted uncertainties are for the 2012 dataset. The uncertainties marked with † are treated as shape uncertainties while the rest are normalisation uncertainties.

The major sources of experimental systematic uncertainties are as follows:

Luminosity

The total integrated luminosity recorded by the ATLAS detector cannot be perfectly measured. The uncertainty on the luminosity measurement is ±2.8% (±1.8%) for 2012 (2011) data, with the method used for the calculation described elsewhere [100]. Any backgrounds modelled by MC are scaled to the recorded integrated luminosity so are affected by this normalisation uncertainty.

Tau Energy Scale

The energy scale uncertainties form a significant group of experimental uncertainties. They derive from the imperfect knowledge of how the detector responds to taus and jets with different properties. They also must account for any mismodelling of tau and jet energies in MC.

The uncertainties lead to variations of the energy of each associated object, meaning they can cause events which previously passed a category selection to now fail and vice versa. This makes the energy scale uncertainties shape uncertainties. As these uncertainties vary the energies of taus and jets, all secondary variables (such as the E_T^{miss}) must be recalculated for each variation.

The tau energy scale (TES) is an important energy scale uncertainty. The method used to calculate the TES differs depending on the tau p_T, meaning this uncertainty has multiple components [49]. A shift is applied to the p_T of taus in data, with the size of the shift determined by examining the $Z \to \tau\tau$ mass peak in data events. The most important TES components relate to the detector model, detector response, underlying event, pile-up and the MC parton shower model.

The combined TES uncertainty is of the order ±2–4%. The tau energy resolution is measured with an uncertainty below ±1%.

Jet Energy Scale and Resolution

The jet energy scale uncertainties (JES) come from a number of components [101]. The impact of each component was carefully evaluated and irrelevant components were discarded. Components were deemed relevant if their impact was $> 10\%$ of that of the leading component. They were also required to have an effect which was distinguishable from statistical fluctuation. This procedure was verified by confirming that the sum in quadrature of the retained components was much larger than that of the discarded components.

The trimming of components is in addition to the pruning and smoothing procedure described above. The procedure was required to ensure statistical noise was not added to the systematic uncertainties, while retaining all necessary information. Furthermore, as each JES component required rerunning the event selection and categorisation, the truncated list guaranteed the analysis could be ran in a timely manner. The JES uncertainties do not apply to the $Z \to \tau\tau$ or fakes background as their modelling uses jets directly from data.

The sources of JES uncertainty are as follows:

- *In-situ jet energy correction*: consists of a statistical component as well as components relating to the detector response and jet modelling in MC
- η *intercalibration*: a statistical component alongside an MC modelling one
- *Jet flavour and composition*: components relating to the quark-gluon composition of the jets as well as how the detector responds differently to light-quark and gluon initiated jets
- *b-jets*: how the detector responds to b-jets
- *Pile-up*: how the underlying event impacts jet energy measurements

Their combined uncertainty is typically around a few percent, increasing to up to $\pm 10\%$ for the VBF signal samples where the JES is of particular relevance.

The jet energy resolution (JER) uncertainty is obtained by smearing the jet energies with a smearing factor related to the JER measurement uncertainty. The smearing produces a one-sided uncertainty and this variation is then symmetrised before use. The JER uncertainty is both a shape and normalisation uncertainty.

The uncertainty on the jet energy resolution measurement is around ± 10–20% [47]. When propagated to the analysis, the effect is of the order of a few percent.

Lepton Energy Scale and Resolution

Similar to taus and jets, leptons are also subject to energy scale and resolution uncertainties.

For electrons, the $Z \to ee$ resonance is used to measure the energy scale and resolution [43]. The energy scale is determined to an accuracy of $\pm 0.05\%$ while the resolution uncertainty is around $\pm 10\%$ for electrons with $E_T < 60$ GeV.

The muon uncertainties are measured using muon pairs from J/ψ, Z and Υ decays [39]. The muon energy scale uncertainty varies in the range ± 0.05–0.2% as a function of η. For a typical muon, the resolution uncertainty is around $\pm 2\%$.

Lepton Trigger Efficiency

Both the electron and muon triggers used in the analysis have associated uncertainties. Tag-and-probe studies are performed to measure any mismodellings of the trigger efficiencies in MC. These studies produce scale factors which are used to reweight the MC samples and these SF have associated uncertainties [38, 102]. The SF and their uncertainties are η and ϕ dependent. The uncertainties are propagated to the analysis as the uncertainty on the lepton trigger efficiencies.

Lepton Reconstruction, Identification and Isolation Efficiencies

Similar to the trigger efficiencies, the lepton reconstruction and identification algorithms have related reweightings with uncertainties [39, 40, 42]. The uncertainties are included in the analysis uncertainty treatment. The lepton uncertainties have a combined impact of around $\pm 2\%$ for the two signal samples.

Tau Identification Efficiency

The hadronic taus are not used for triggering but do have efficiency uncertainties due to the reconstruction and identification procedures [49]. The relevant SF and uncertainties are determined by performing a tag-and-probe study on $Z \to \tau\tau$ events. The SF and uncertainties are only applied to tau candidates which are matched to a true tau. The true particles in an MC event are recorded, before detector simulation. A tau is considered truth matched if a true tau lies within $\Delta R < 0.2$ of the tau candidate.

Alongside the tau identification SF, correction factors are also required for the $e \to \tau$ fake rate in MC. These SF and their uncertainties are only applied to tau candidates truth matched to an electron.

The tau identification uncertainties are around $\pm 3\%$ for the signal samples and ± 1–2% for the background.

Missing Transverse Energy

The E_T^{miss} is recalculated for each systematic variation which impacts the energy measurements of the taus, jets and leptons. In addition, the E_T^{miss} has two dedicated uncertainties.

The E_T^{miss} of each event is calculated by summing the energy of each object in the event and taking the imbalance. There is an additional contribution which originates from the objects which are too low-p_T to be fully reconstructed. This term is named the "soft term" and the E_T^{miss} uncertainties relate to the scale and

resolution of the soft term. Two methods are used to calculate the soft term's scale and resolution uncertainties [50].

The first method uses the data-MC ratio in $Z \to \mu\mu$ events with zero jets. Such events are chosen as only the muons and the soft term contribute to the E_T^{miss} calculation. The E_T^{miss} is projected onto the direction of the Z^0 boson, to check for any bias in its soft term. The mean deviation from unity of the data-MC ratio for this projection is used to determine the scale uncertainty. Any incompatibility with the resolution of the x- and y-components of the E_T^{miss} is taken as the resolution uncertainty of the soft term.

The second method uses the imbalance between the E_T^{miss} and the p_T sum of the hard objects to measure the uncertainties. Electrons, muons, jets, photons and neutrinos are the hard objects used to construct the p_T sum. The use of neutrinos means this method can only be applied to MC, as the neutrino kinematics cannot be accurately determined in data. An inclusive $Z \to \mu\mu$ sample is used for this method.

The expectation is that the soft term and hard-p_T sum are equal in size and in opposite directions. The soft term is decomposed into two components, one collinear and one orthogonal to the hard-p_T sum. The scale and uncertainty of the soft term can be measured using the collinear component. The perpendicular term can be used to determine the soft term resolution and uncertainty.

Both methods give consistent results for the uncertainties. The combined impact of the two soft term uncertainties on the overall E_T^{miss} is below $\pm 1\%$.

b-Tagging

The use of b-tagging in the analysis leads to additional uncertainties. These again relate to correction factors due to MC mismodellings and their associated uncertainties. The b-tagging uncertainties are of the order of $\pm 2\%$ for jets with $p_T < 100$ GeV [46].

8.2.3 Background Modelling Uncertainties

Despite the aforementioned simplification of the experimental systematic uncertainties when using data driven techniques, additional uncertainties are required. These are termed background modelling uncertainties and exist for $Z \to \tau\tau$, modelled by the embedding method, and the fakes background, estimated using the fake factor method.

The embedding technique has two major associated systematic uncertain-

ties. These arise from the selection of the $Z \to \mu\mu$ events and the calorimeter cell subtraction, described in Section 6.3. The uncertainty on the $Z \to \mu\mu$ event selection is estimated by both tightening and removing the isolation requirements on the two muons. The calorimeter cell subtraction systematic is estimated by varying the energy subtraction by $\pm 20\%$ ($\pm 30\%$) for the 2012 (2011) embedding samples. These two variations should account for any bias introduced by the chosen isolation criteria or the calorimeter energy subtraction procedure.

In addition to the two major systematics, the embedding samples also have a number of less significant modelling uncertainties. Weight corrections are applied to the $Z \to \mu\mu$ events to account for muon triggering and reconstruction variations. These corrections have associated uncertainties which are propagated through to the final embedding sample.

The fake factor method has four major associated uncertainties. The first such uncertainty is simply the statistical uncertainty from the fake factor calculations. The inputs to the FF calculations each have a statistical uncertainty and this is propagated through to the stage where the FF are applied to data events.

The mixing ratios, R_i, are also subject to some uncertainty. A systematic uncertainty is defined by varying each of the ratios up and down by a factor of two and examining the effect on the fake estimation. This factor of two, while somewhat arbitrary, was determined to be a suitably conservative approach. The variation is found to cause a $\pm 5\%$ change for the 2012 sample and a ± 10–15% effect for the 2011 sample (dependent on the signal category). These numbers are retained for use as the systematic uncertainty.

The third and fourth fake factor uncertainties originate from closure tests performed on the fake factor method. For the Z+jets, W+jets and top backgrounds, MC samples are available. These were used to calculate a new set of FF and these FF were then applied to the MC samples from which they were derived. As the FF are tau p_T dependent, this results in perfect agreement when plotting the tau p_T spectrum in the signal region using the MC samples. However, some bias may be introduced in the BDT score. No statistically significant effect is observed for the 2012 samples whereas a 10% deviation is present for 2011. This 10% is taken as the third systematic uncertainty for the 2011 fake factors.

Fourthly and finally, a second closure test is required as the QCD multijet background has no available MC samples. Any deviation between the observed data and the fakes prediction in the SS data CR is taken as this systematic uncertainty. Again, no significant discrepancy is present for the 2012 samples however a $\pm 10\%$ systematic uncertainty is required for the 2011 data.

8.2.4 Theoretical Uncertainties

A range of theoretical uncertainties are considered in this analysis, primarily relating to the MC signal samples used in the analysis. Simulated background samples are also subject to theoretical uncertainties whereas the data driven background models are not.

All signal samples have uncertainties due to higher-order QCD corrections not included in their cross section calculations. These are estimated by varying the factorisation and renormalisation scales up and down by a factor of two. Before calculating these uncertainties, category appropriate parton-level cuts were applied to the samples. These come in the form of a cut on the jet kinematics for the VBF category and on the Higgs p_T for the boosted category. For the ggF samples, the uncertainties are approximately $\pm 24\%$ for both categories. They are significantly lower for the VBF and VH samples, in the range of ± 2–4%.

No veto is placed on the presence of a third jet in the VBF category, however the trained BDT strongly disfavours the presence of a third jet in the event. As the ggF+3 jets cross section is only known at LO, this leads to a large uncertainty on the ggF contamination in the high BDT score bins for the VBF category. The ggF uncertainty on the highest score bin is $\pm 30\%$, although the bin is dominated by VBF events, somewhat reducing the impact of this uncertainty in the final fit.

A comparison is made between signal samples generated with Powheg+Pythia and Powheg+Herwig. Pythia and Herwig are two parton shower simulators which use different methods to model multi-parton interactions. This makes the comparison a method for determining the uncertainty due to the underlying event. The test is considered for the ggF and VBF signal samples as the VH contribution is found to be insignificant. The uncertainty is calculated to be in the range of ± 1–10%, depending on the category and signal process. No significant BDT shape variation is observed.

The uncertainty due to the choice of PDF is also relevant for the ggF and VBF categories. It is calculated by reweighting signal samples produced using the default CT10 NLO PDF set to the MSTW2008NLO [103], NNPDF2.1 NLO [104] and CT10 eigen-tune parameterisation predictions. The uncertainty is taken as the largest difference in acceptance between the nominal PDF and the reweightings. A shape dependent uncertainty is found, with variations of the order $\pm 5\%$ for the ggF signal sample and $\pm 1\%$ for VBF. The bin-by-bin uncertainty is shown in Figure 8.1.

Different MC generators were also tested for significant differences in acceptance. In each case, the two generators which are compared use different methods for the matrix element calculation and the matching between the matrix element and

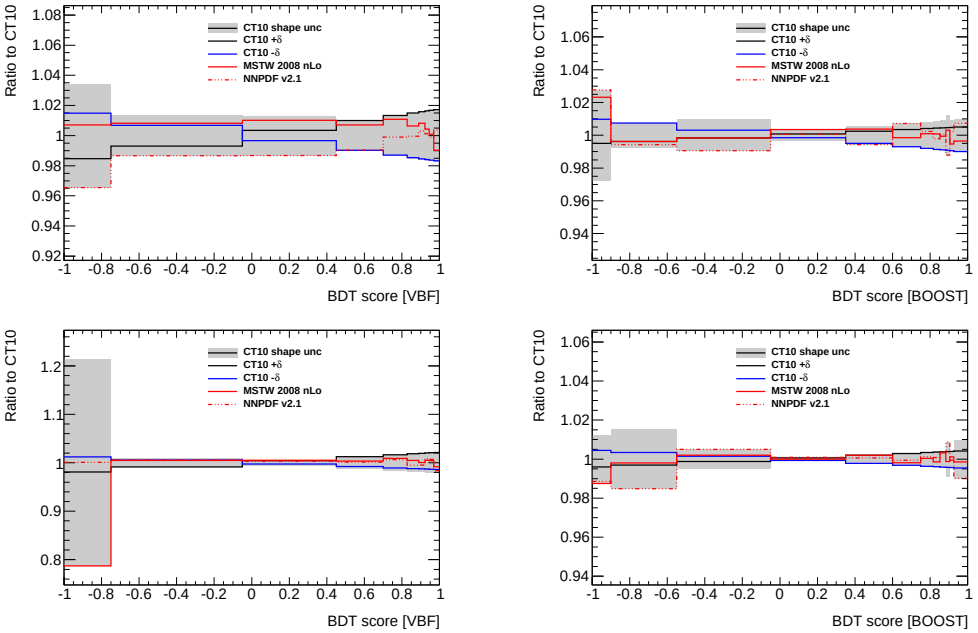

Figure 8.1: The bin-by-bin PDF shape uncertainty, split by category and signal sample. The left-hand plots are for the VBF category and the right are for the boosted category. The top row is for the VBF signal sample with the bottom row showing the ggF sample. The black and blue lines show the $\pm 1\sigma$ variations of the CT10 eigen-tune parameterisation. The solid and dashed red lines show the impact of the MSTW2008NLO and NNPDF2.1 NLO reweightings respectively. The grey band shows the final shape uncertainty, defined as the symmetrised maximum variation for the bin.

Category	BDT Bin Edges
VBF	$-1.0, -0.95, -0.35, 0.35, 0.7, 0.851, 0.904,$ $0.936, 0.955, 0.969, 0.979, 0.988, 1.0$
Boosted	$-1.0, -0.95, -0.75, -0.4, 0.0, 0.307, 0.56,$ $0.704, 0.804, 0.855, 0.904, 1.0$

Table 8.2: Bin edges of the two analysis categories used in the analysis fit, as determined by the dedicated binning algorithm.

parton showering. For ggF, the Powheg+Herwig sample was compared with an MC@NLO+Herwig [105] one. The VBF Powheg+Herwig sample was compared to an aMC@NLO+Herwig [106] sample. The acceptance uncertainty is found to be in the range ±2–4%, varying by category and signal sample.

The final signal uncertainty considered is the uncertainty on the $H \to \tau\tau$ branching ratio. This is calculated to be ±5.7% [78].

8.3 BDT Binning Studies

The choice of binning for the BDT score when performing the fit has a significant impact on the fit sensitivity. Finer binning allows the separation between signal and background expectations to be exploited. However, this must be balanced with the increased statistical uncertainty which finer binning introduces. It is also desirable that each bin has a robust background prediction. An algorithm which satisfies these requirements was developed and used to determine the binning used in the fit.

The algorithm starts with the BDT score distribution finely split into equal width bins. Beginning from the high-score bins on the right of the BDT score histogram, bins are merged until the relative statistical uncertainty on each background is below a specified threshold. The thresholds are set differently for each background, depending on the background's perceived importance. They range from a low of 30% for the $Z \to \tau\tau$ background, up to 100% for top processes. It is also required that each bin has an increase in the predicted number of background events when compared to the bin on its right. The bins defined by this algorithm are shown in Table 8.2 and are used for the global fit.

8.4 Fit Model Checks

Confirming the fitting procedure behaves in a reasonable manner is an important step in validating the results of the fit. The pulls and constraints on each NP are

evaluated individually. Furthermore, the impact of each NP on the best-fit μ is determined and the NPs are ranked in order.

Pre-fit, each NP is assigned a nominal (central) value and an uncertainty. These values form an input to the fit, alongside the event counts in the signal BDT distributions. Following the convergence of the global fit, each NP has a post-fit central value and associated uncertainty. Generally, in a well behaved fit, the pre-fit and post-fit values should be compatible.

An NP is said to be pulled if the pre- and post-fit central values are significantly different. This is usually expressed by taking the difference between the pre- and post-fit central value and dividing it by the pre-fit uncertainty. When expressed in this way, the pull should be around zero. However, some fluctuations about zero are expected, especially considering the large list of NPs considered in this analysis.

In a similar vein, the constraint on an NP is measured as the ratio between the pre- and post-fit uncertainty. This value should be of order one, but again, some fluctuations are expected.

It is, however, possible for the post-fit values to significantly differ from the pre-fit ones. Such a large difference would suggest the NP does not accurately describe the uncertainty and would demand further study. In such cases it is important to fully understand where the data's power to pull or constrain originates from.

The ranking of the NPs is determined by how strongly each NP affects the fitted value of μ. In turn, each NP is fixed to its post-fit value plus or minus its post-fit uncertainty. The global fit is then re-run, with all but the fixed parameter allowed to vary as usual. The parameter ranking is calculated as the change in best-fit μ when fitting with the NP held at its extreme value, divided by the total uncertainty on μ. For example, for the simple case of a fit with just one NP, that NP would score one in this test. This is because the single NP would be the source of 100% of the uncertainty on μ.

While the ranking of NPs alone cannot indicate problems with the fit, it can help uncover bugs in the systematic uncertainty implementation. For example, many NPs are not expected to have a large impact on the results. The appearance of such NPs high in the ranking would be a cause for further investigation.

Figures 8.2, 8.3 and 8.4 show the rankings and pulls for the VBF category, boosted category and their combination. The rankings of the NPs appears logical from a physical point of view. For example, statistical fluctuations in the bins with a high BDT score has the leading impact on μ in both categories. In the combination, the jet energy scale uncertainties score highly but this is to be expected. Some minor pulls are observed but all post-fit values are within $\pm 1\sigma$ of their pre-fit value.

These factors combined suggest the global fit of the analysis is well behaved.

In the full $H \to \tau\tau$ analysis (as opposed to the $H \to \tau_\ell \tau_{had}$ specific case discussed here), the three subchannel fits can be performed together [52]. Common NPs in the three subchannels may be correlated, depending on the underlying physics. The additional information provided by the other decay channels is used to simultaneously constrain the NPs in the fit.

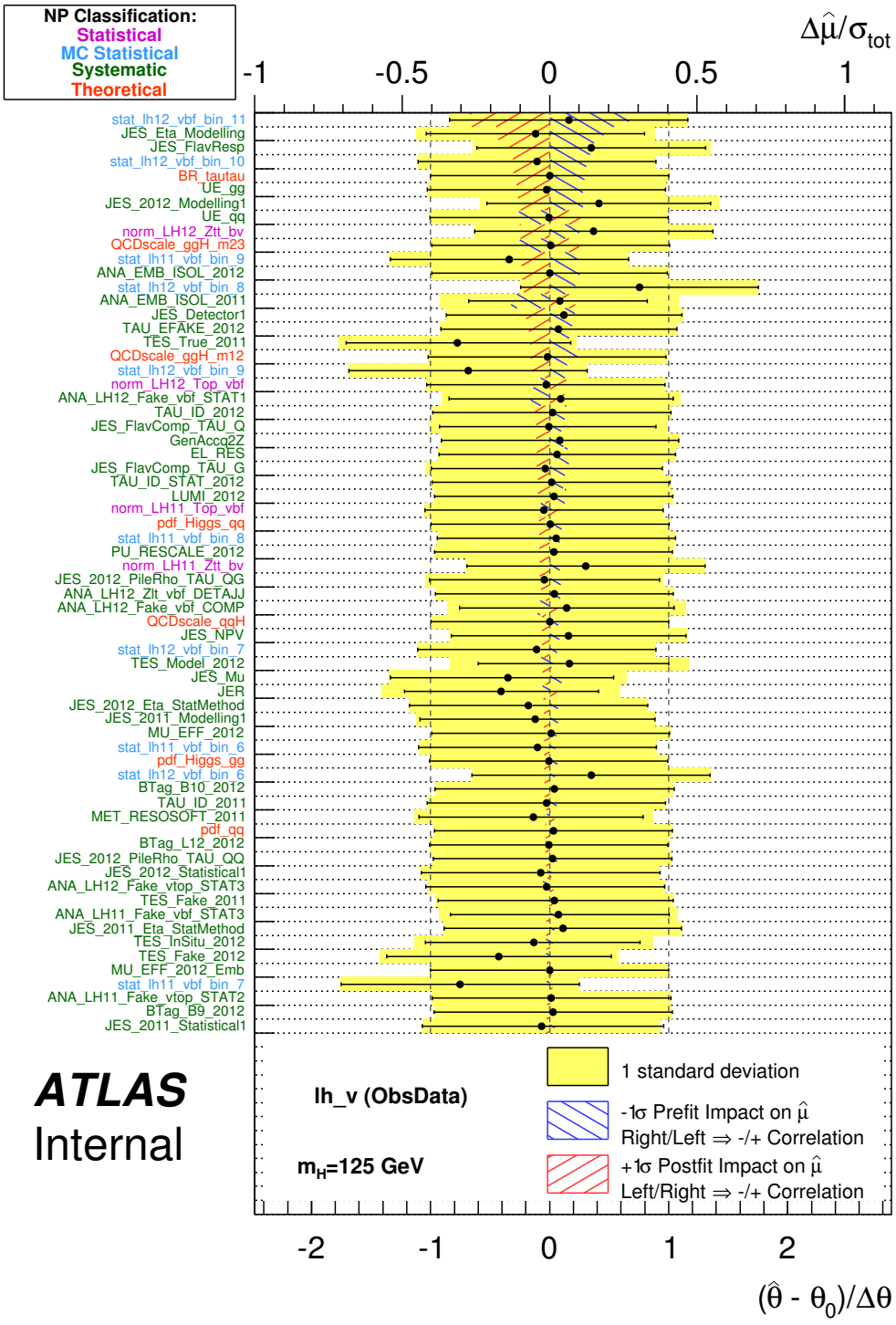

Figure 8.2: Nuisance parameter rankings and pulls for the VBF category [52]. Each NP is briefly described in Appendix B.

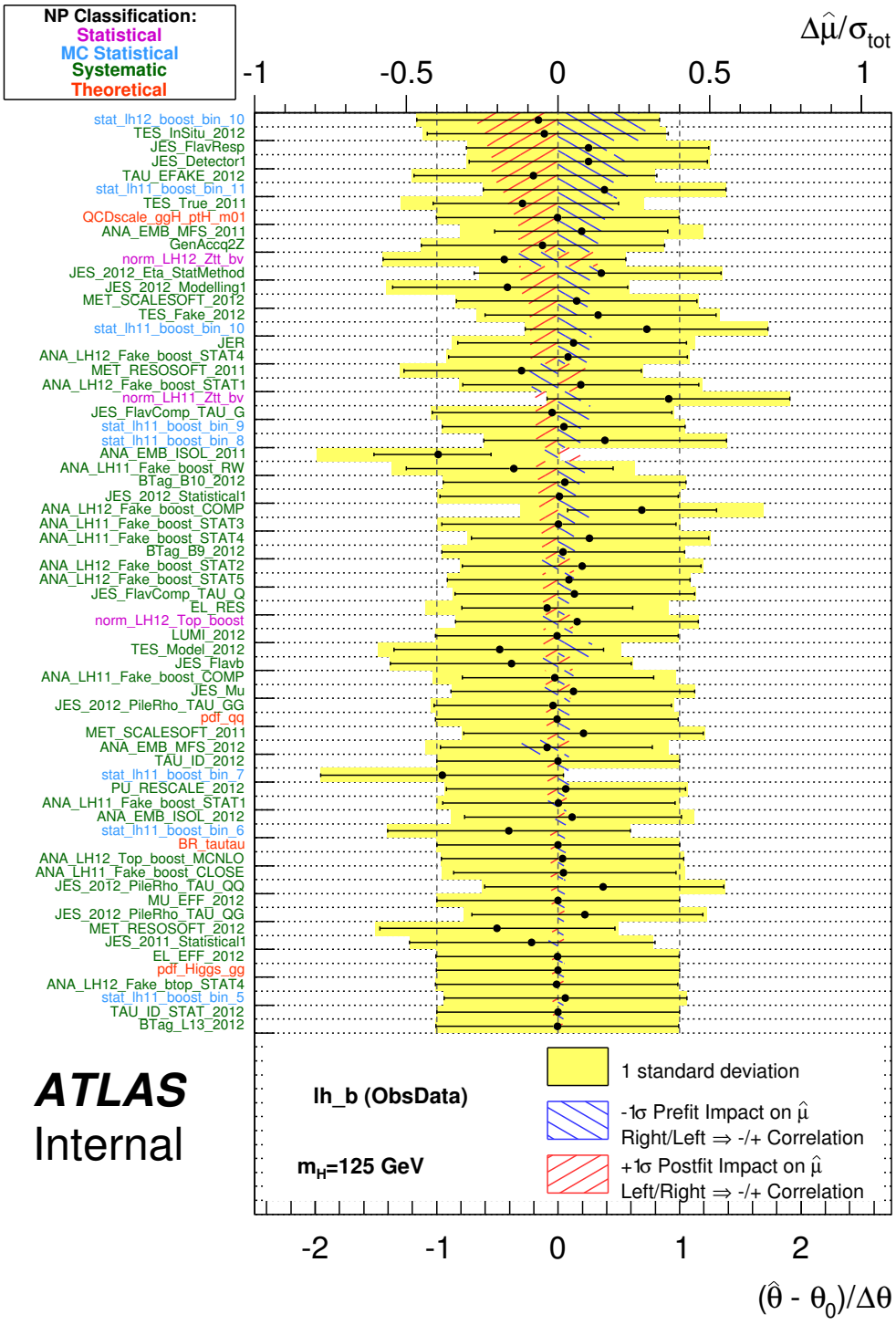

Figure 8.3: Nuisance parameter rankings and pulls for the boosted category [52]. Each NP is briefly described in Appendix B.

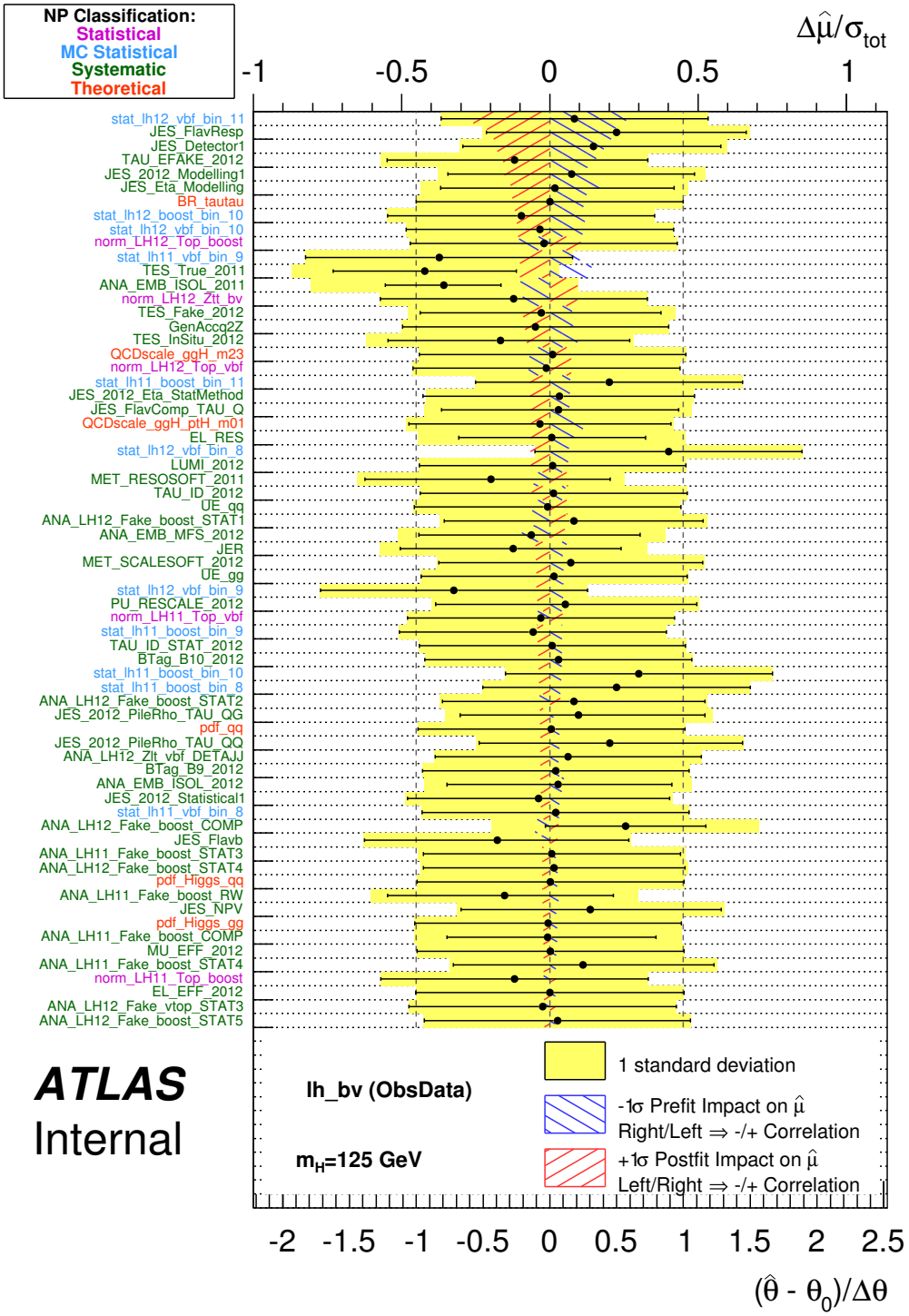

Figure 8.4: Nuisance parameter rankings and pulls for the combination of the VBF and boosted categories [52]. Each NP is briefly described in Appendix B.

Chapter 9

Results and Conclusions

9.1 $H \to \tau_\ell \tau_{had}$ Results

The signal strength parameter, μ, is evaluated assuming a Higgs mass of $m_H = 125.36$ GeV. This is the best-fit value of the Higgs mass, as measured by the combined $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ analyses of the ATLAS collaboration [9]. The signal strength parameter is quoted in units of the SM cross section, with $\mu = 1$ equivalent to an observation of the Higgs with the production cross section predicted by the Standard Model. The SM cross sections used are given in Table 6.1.

The fit is a maximum likelihood fit of the BDT score distributions, as described in Chapter 8. For the $H \to \tau_\ell \tau_{had}$ analysis, the signal strength parameter is measured to be

$$\mu = 1.0^{+0.4}_{-0.3}(\text{stat.})^{+0.4}_{-0.3}(\text{syst.})^{+0.1}_{-0.1}(\text{theory}). \tag{9.1}$$

Here, the uncertainty labelled "theory" refers only to the combined uncertainties on the Higgs production cross section and the $H \to \tau\tau$ branching ratio. Any other theory uncertainties are included in the quoted "syst." uncertainty. The systematics are discussed in detail in Section 8.2 and their individual impacts are summarised in Table 8.1.

Using Equation 8.2 and setting $\mu = 0$ allows the probability that the observation is due to the background only hypothesis to be calculated. This value is known as the p-value and is approximately 1% for the observed result. This corresponds to a 2.3σ observation of an excess.

The $H \to \tau_\ell \tau_{had}$ result can be subdivided into signal strength measurements in the two analysis categories. The VBF channel has a best-fit μ value of $1.0^{+0.6}_{-0.5}$. For the boosted channel the result is $\mu = 0.9^{+1.0}_{-0.9}$. The post-fit BDT distributions for the VBF and boosted categories are shown in Figure 9.1.

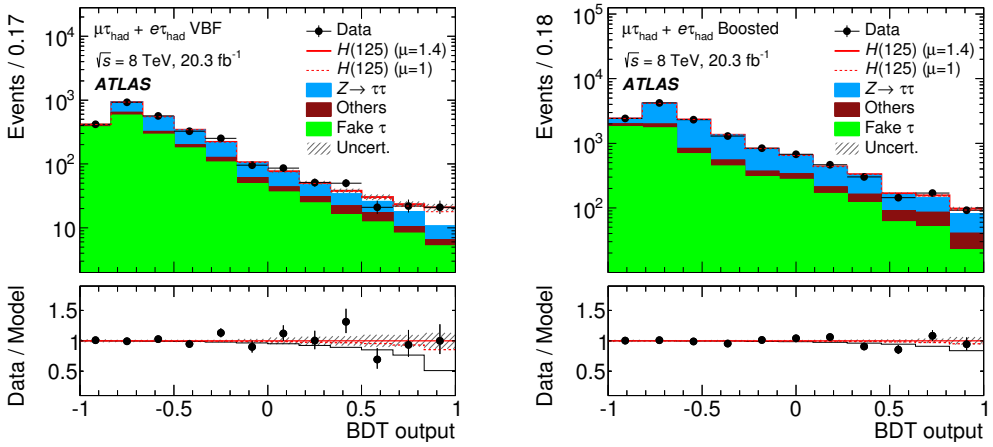

Figure 9.1: Post-fit BDT distributions for the VBF (left) and boosted (right) categories [52]. The signal expectation is shown for both the SM ($\mu = 1$) and combined best-fit ($\mu = 1.4$) hypotheses.

9.2 $H \to \tau\tau$ Combination

As well as the $H \to \tau_\ell \tau_{had}$ analysis, parallel analyses are performed which search for $H \to \tau\tau$ decays where both taus decay leptonically or both decay hadronically. These analyses share many features with the $H \to \tau_\ell \tau_{had}$ analysis but also differ in some respects. Shared characteristics include the use of BDTs in two categories which individually target VBF and boosted events. The sub-channels differ in some of their object definitions, background modelling techniques and BDT input variable choices.

Combining the results of the three sub-channels gives

$$\mu = 1.43^{+0.27}_{-0.26}(\text{stat.})^{+0.32}_{-0.25}(\text{syst.})^{+0.09}_{-0.09}(\text{theory}). \tag{9.2}$$

The combined results have a p-value of around 3×10^{-6} which corresponds to a 4.5σ observation. The expected sensitivity of the analysis is 3.4σ. Figure 9.2 summarises the results, broken down by year, category and sub-channel. No significant deviations from the Standard Model (SM) expectation are observed.

To better visualise the results of the fit, a plot was produced which combines the BDT distributions from each sub-channel and category which were used in the fit. Each bin is included in the plot and the bins are ordered by their $\log_{10}(S/B)$ (where S and B are the expected number of signal and background events respectively). This results in a distribution which is background dominated on the left and more signal-rich on the right. The plot is shown in Figure 9.3.

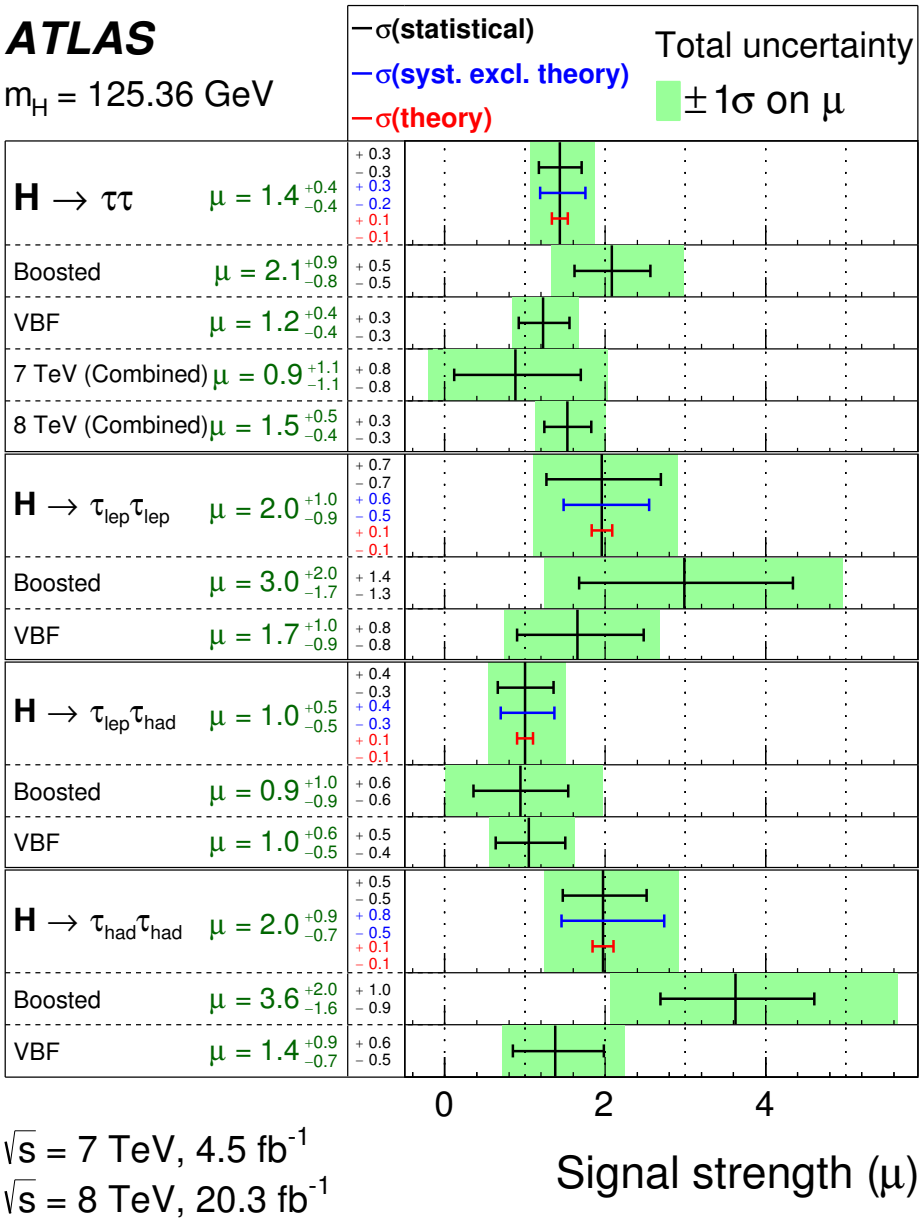

Figure 9.2: Best fit signal strength parameter measurements, broken down by year, category and sub-channel [52]. The green bands represent $\pm 1\sigma$ variations of the total uncertainty. Uncertainties are also shown separated into statistical (black), systematic (blue) and theoretical (red) components.

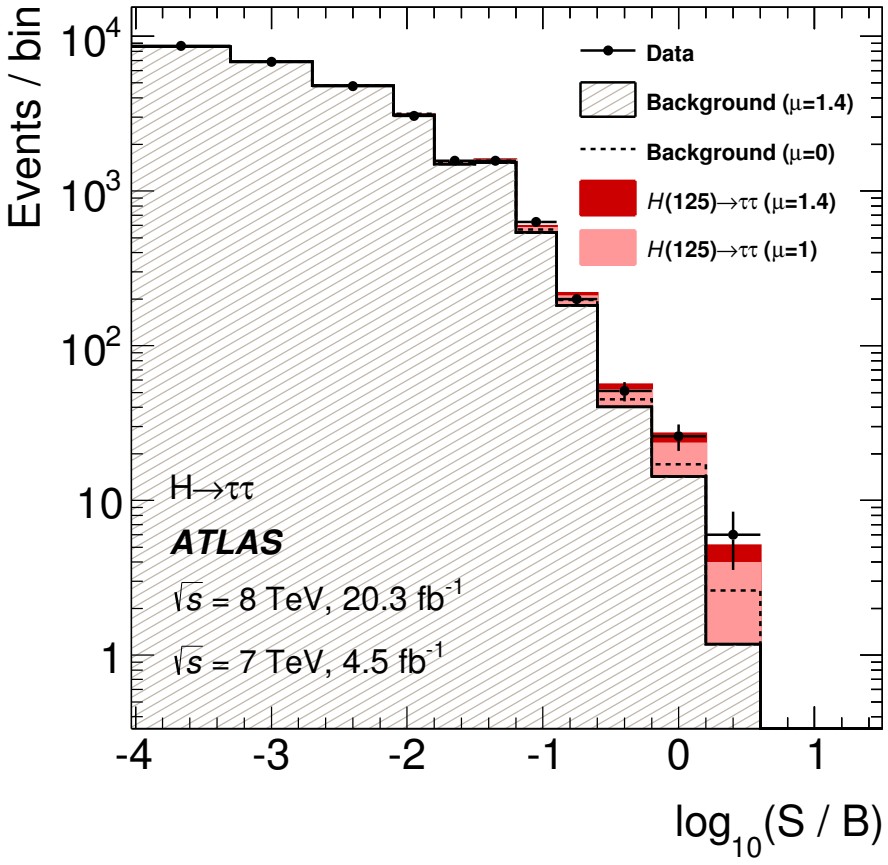

Figure 9.3: All bins used in the combined fit of the analysis, ordered by $log_{10}(S/B)$ (where S and B are the expected number of signal and background events respectively) [52]. The background prediction is displayed for both the best-fit ($\mu = 1.4$) and background-only ($\mu = 0$) hypotheses. The signal prediction is displayed for both the best-fit and SM ($\mu = 1$) expectations.

The signal hypothesis is included for both the SM expectation ($\mu = 1$) and the best-fit signal strength ($\mu = 1.4$). The background hypothesis is displayed for the $\mu = 1.4$ case, with the background expectation in the signal-free scenario also included as a dashed line. The leftmost bins show good agreement, demonstrating that the background processes are well modelled. The rightmost bins show the data is compatible with the SM expectation and is incompatible with the background-only hypothesis.

A weighted mass plot is also produced, where each event used in the fit is weighted by the $\ln(1 + S/B)$ of its BDT bin. This weighting serves to enhance events which are most compatible with the signal hypothesis. Although the use of $m_{\tau\tau}^{\mathrm{MMC}}$ in the BDT leads to events with ditau masses around 125 GeV being emphasised, the aim of the plot is to test whether the data is most compatible with the 125 GeV signal hypothesis. The plot is shown in Figure 9.4.

The lower part of Figure 9.4 shows the difference between the weighted data events and weighted background events compared to the weighted signal expectation for three different mass hypotheses, $m_H = 110$, 125 and 150 GeV. Each prediction is plotted using the best-fit signal strength for that mass hypothesis. It can be seen that the data is most compatible with the $m_H = 125$ GeV hypothesis, when compared to the two alternatives presented.

The combined fit is repeated, using two independent signal strength parameters which distinguish between vector-boson-mediated (VBF and VH) and gluon-mediated (ggF) production mechanisms. This approach found

$$\begin{aligned} \mu_{\mathrm{ggF}} &= 2.0^{+0.8}_{-0.8}(\mathrm{stat.})^{+1.2}_{-0.8}(\mathrm{syst.})^{+0.3}_{-0.3}(\mathrm{theory}), \\ \mu_{\mathrm{VBF}+VH} &= 1.24^{+0.49}_{-0.45}(\mathrm{stat.})^{+0.31}_{-0.29}(\mathrm{syst.})^{+0.08}_{-0.08}(\mathrm{theory}). \end{aligned} \tag{9.3}$$

These results are summarised in Figure 9.5, including 68 and 95% confidence limits. Again, the results are fully compatible with the SM expectation.

For the gluon-mediated processes, the observed (expected) significance of the result is 1.74σ (0.95σ). For vector-boson-mediated processes, the significance is 2.25σ (1.72σ).

9.3 Conclusion

A search for $H \to \tau_\ell \tau_{had}$ has been presented, utilising boosted decision trees with the full Run 1 ATLAS dataset. For $m_H = 125.36$ GeV, a 2.3σ excess was observed, with a best fit signal strength of $\mu = 1.0 \pm 0.5$. A combined result, incorporating the $H \to \tau_\ell \tau_\ell$ and $H \to \tau_{had}\tau_{had}$ subchannels, was also presented. There, a 4.5σ

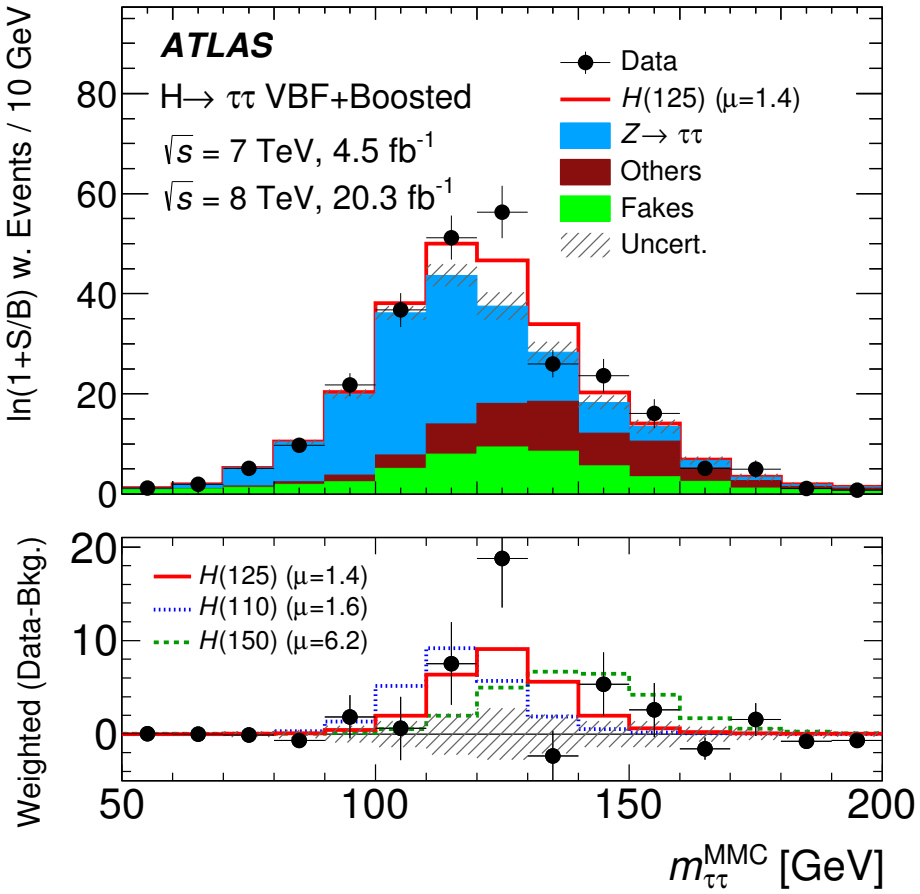

Figure 9.4: Mass plot containing every event used in the fit, weighted by the $\ln(1 + S/B)$ of its BDT bin [52]. This enhances events most compatible with the signal hypothesis. Data is shown using black points. The lower plot shows the difference between the weighted data events and weighted background events (black points). The best-fit signal hypothesis is included for $m_H = 110$, 125 and 150 GeV for comparison.

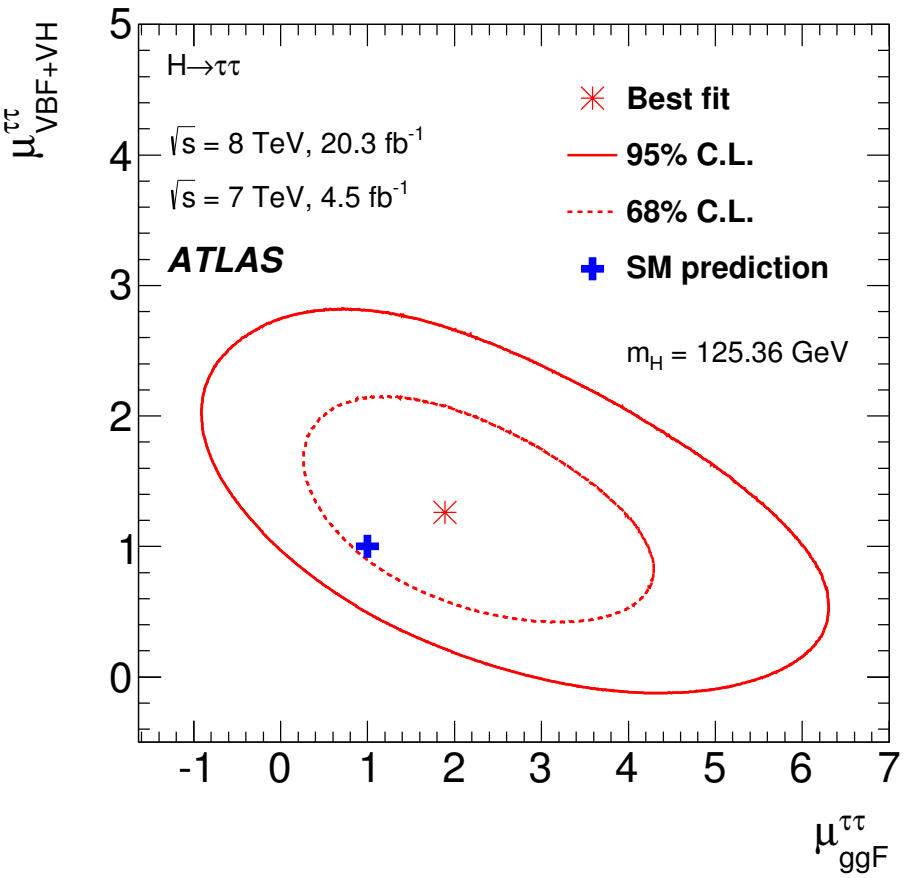

Figure 9.5: Best-fit signal strength parameters, when fitting separately for production processes mediated by vector bosons (VBF and VH) and gluons (ggF) [52]. 68 and 95% confidence limits are included, as are the Standard Model predictions.

excess was measured, giving a signal strength of $\mu = 1.4 \pm 0.4$.

This result constitutes the first direct evidence for the Higgs boson coupling to leptons via the Yukawa coupling. The result is compatible with Standard Model predictions, however the 30% uncertainty means it is too early to draw conclusions. Unlike other decay channels where the Higgs boson has been observed, the most significant contribution to the result is from the VBF Higgs production mechanism.

The $H \to \tau\tau$ era of precision measurement is now being entered. During Run 2 of the LHC, the ATLAS experiment expects to record 300 fb^{-1} of data. This is over a ten times increase on the data recorded during Run 1.

During Run 2, a 5σ observation of $H \to \tau\tau$ should be achievable. Measurements of the Higgs properties in the $H \to \tau\tau$ channel will also become possible. These properties include the mass of the Higgs as well as its spin, parity and charge conjugation quantum numbers. Although these properties have been measured in other Higgs decay channels, $H \to \tau\tau$ is likely to be the first fermionic decay channel capable of such measurements.

Appendix A

Muon Trigger Scale Factors

A.1 $\eta - p_T$ Scale Factors

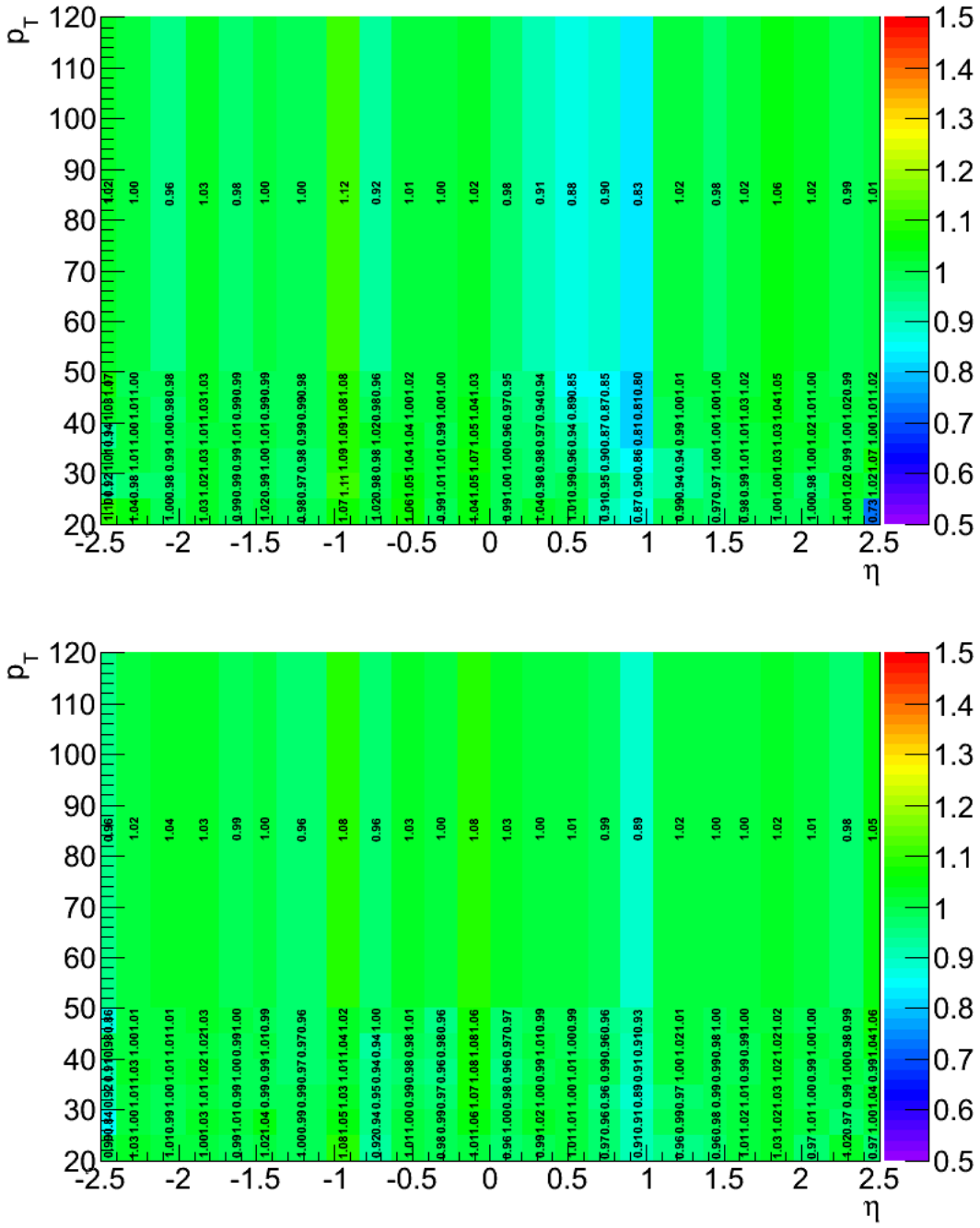

Figure A.1: Derived $\eta - p_T$ SF for periods B–G5, given separately for positively (top) and negatively (bottom) charged muons.

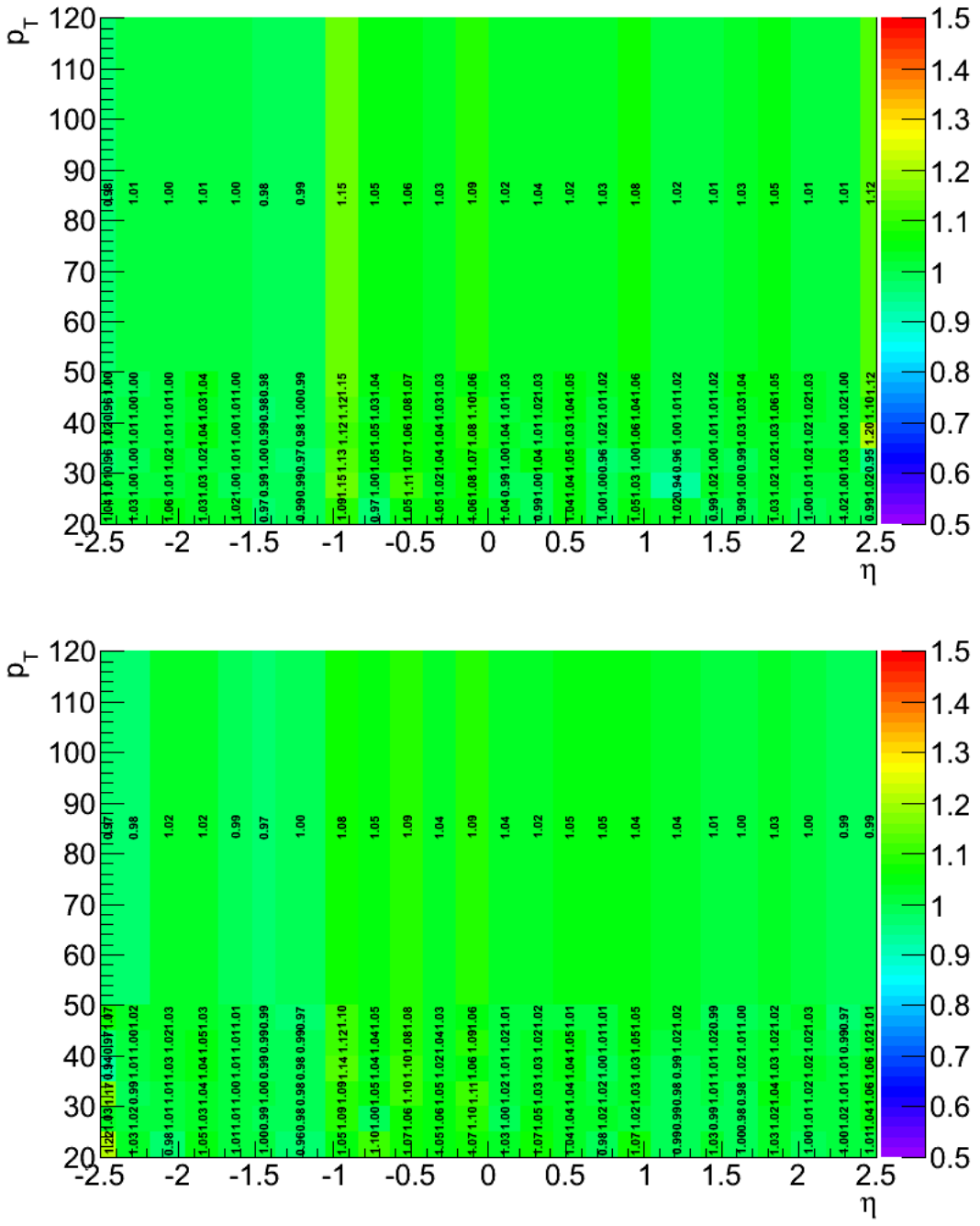

Figure A.2: Derived $\eta - p_T$ SF for periods G6–I, given separately for positively (top) and negatively (bottom) charged muons.

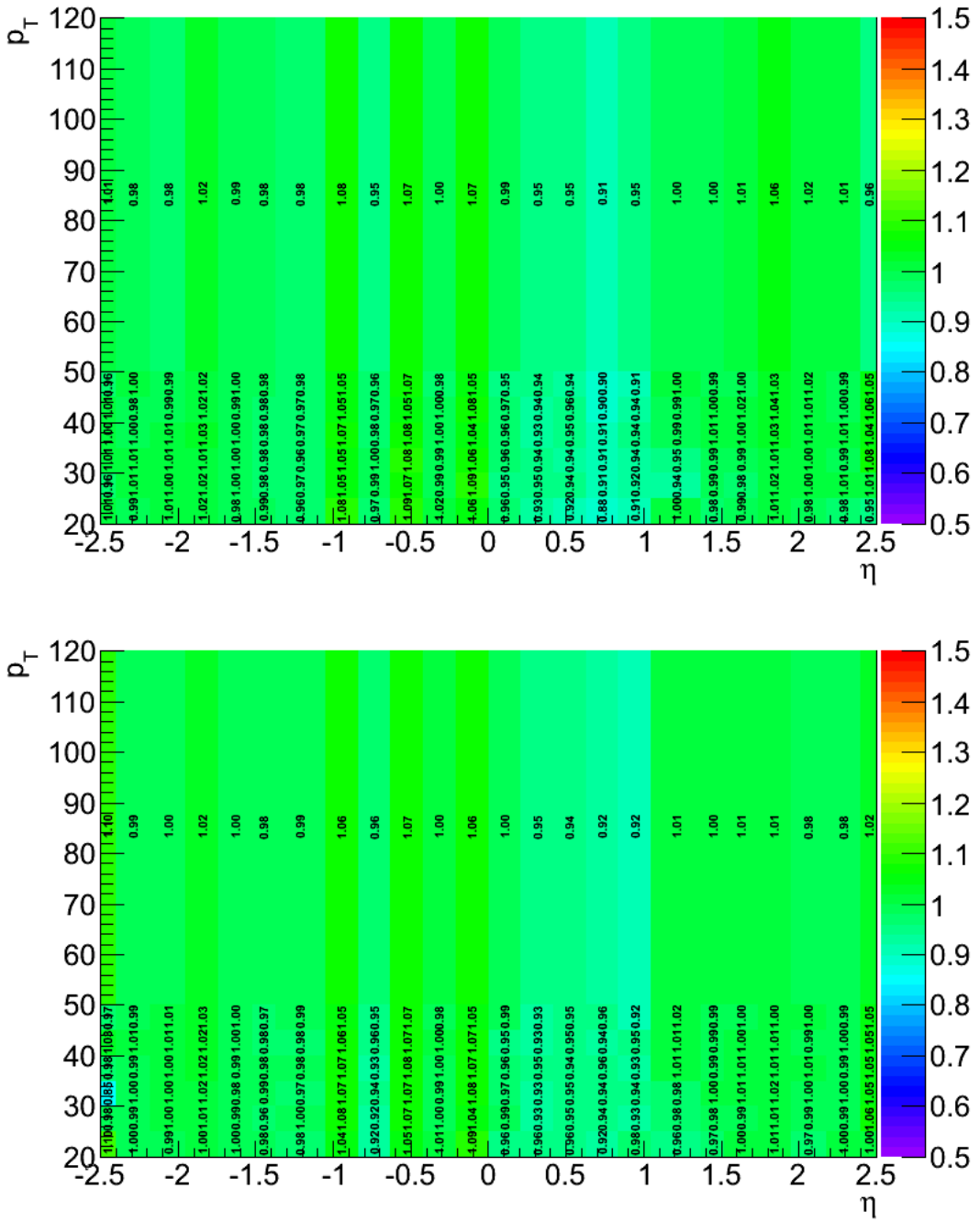

Figure A.3: Derived $\eta - p_T$ SF for periods J–M (excl. L3 and L4), given separately for positively (top) and negatively (bottom) charged muons.

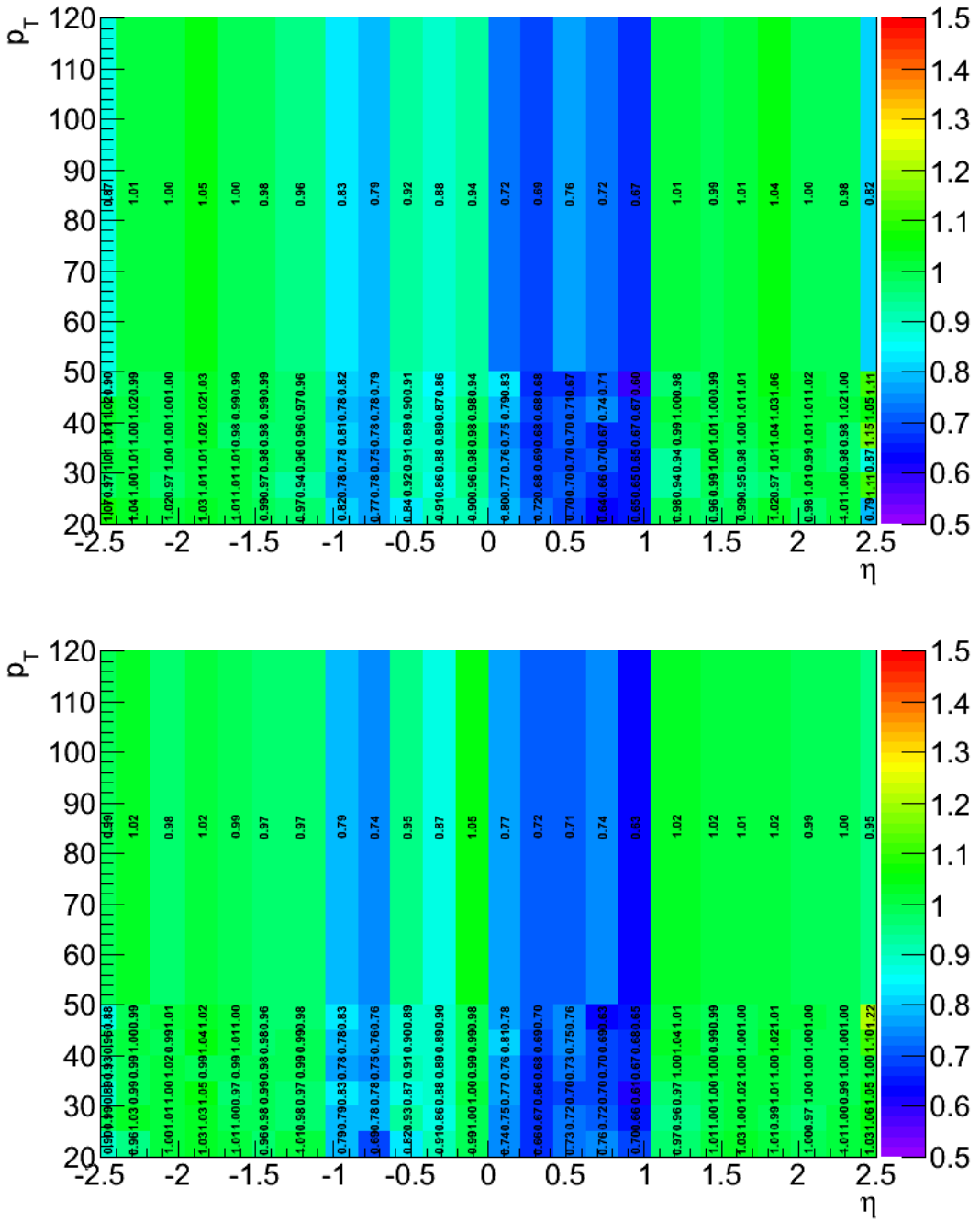

Figure A.4: Derived $\eta - p_T$ SF for periods L3 and L4, given separately for positively (top) and negatively (bottom) charged muons.

A.2 $\eta - \phi$ Second Order Scale Factors

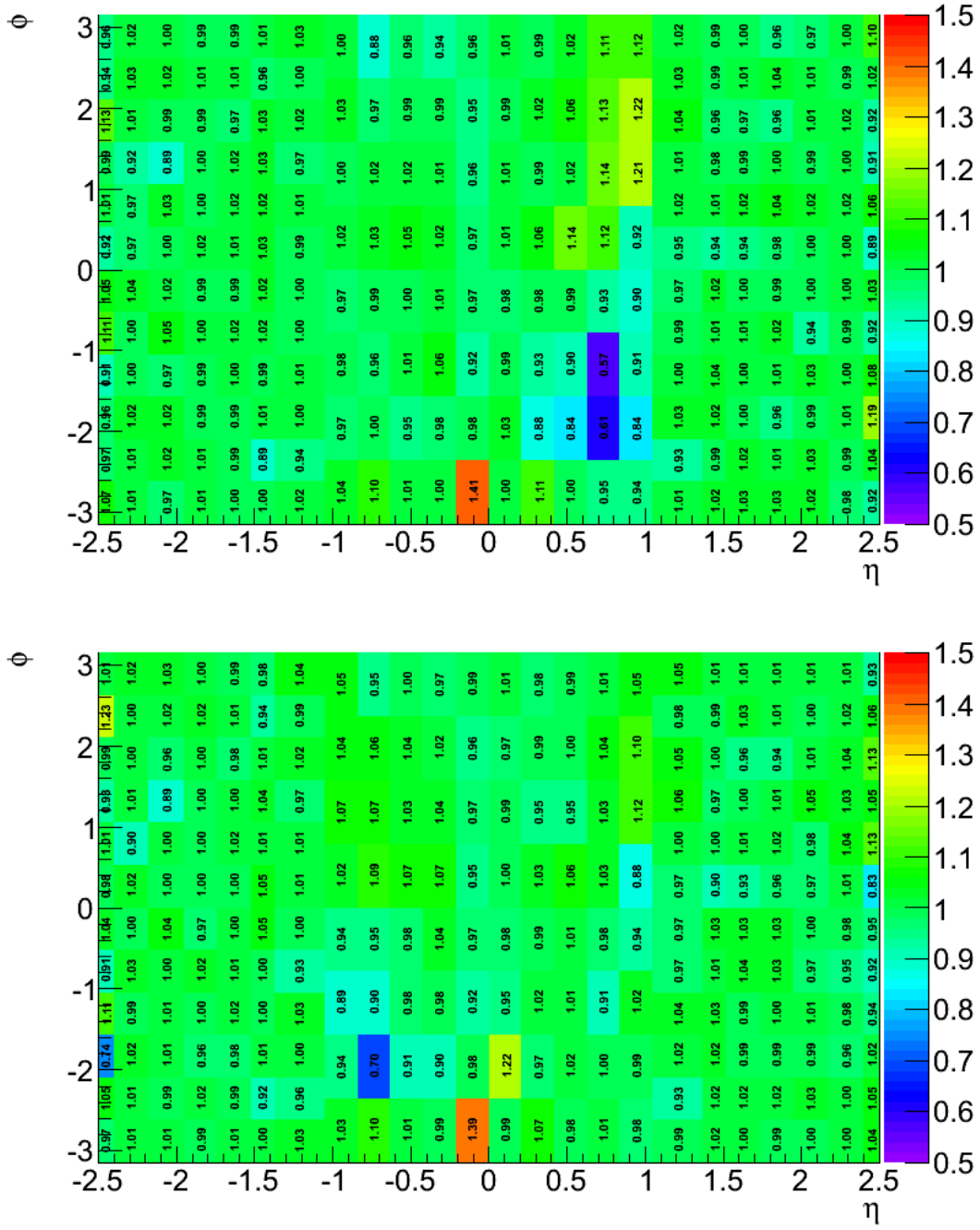

Figure A.5: Derived $\eta - \phi$ SF for periods B–G5, given separately for positively (top) and negatively (bottom) charged muons.

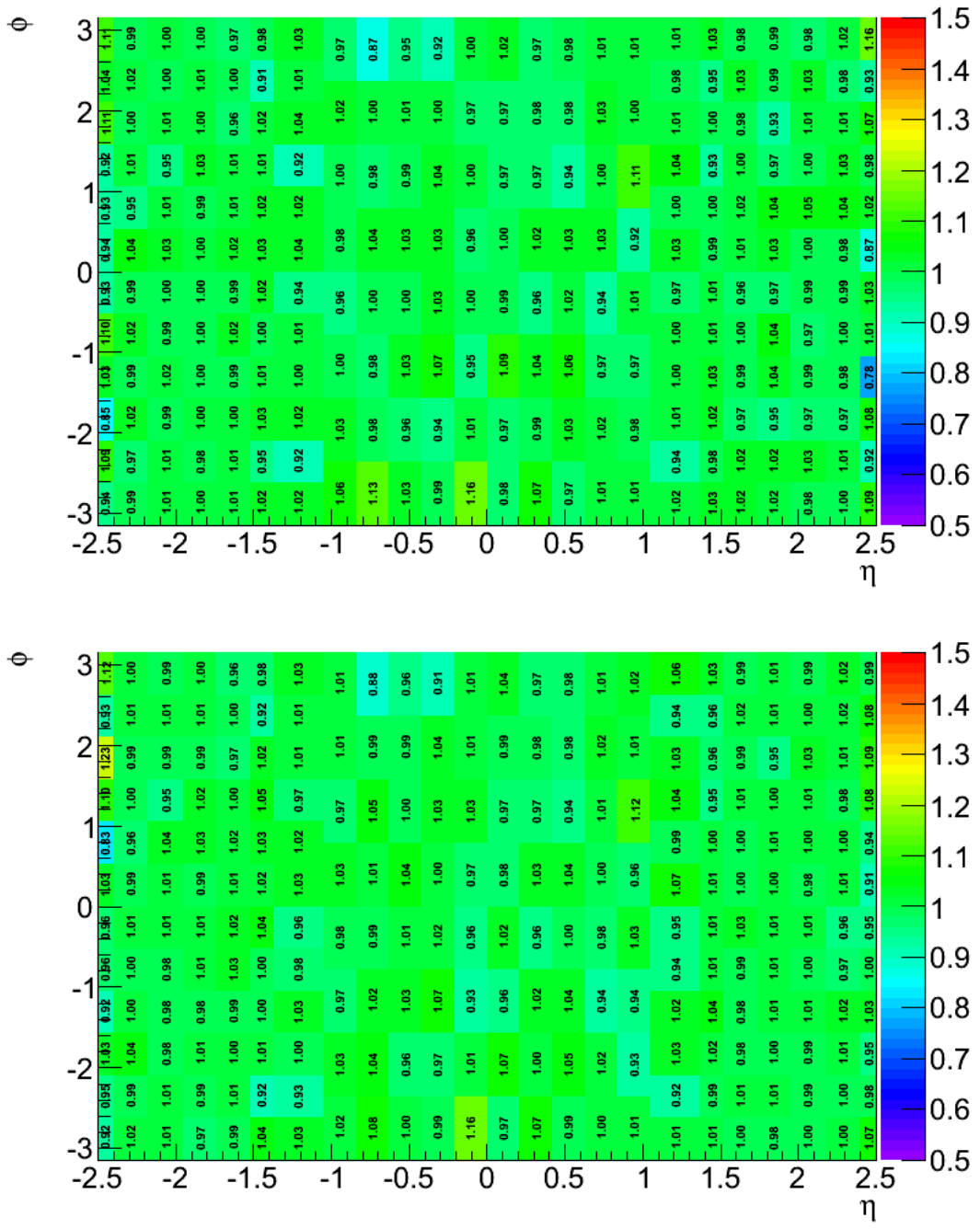

Figure A.6: Derived $\eta - \phi$ SF for periods G6–I, given separately for positively (top) and negatively (bottom) charged muons.

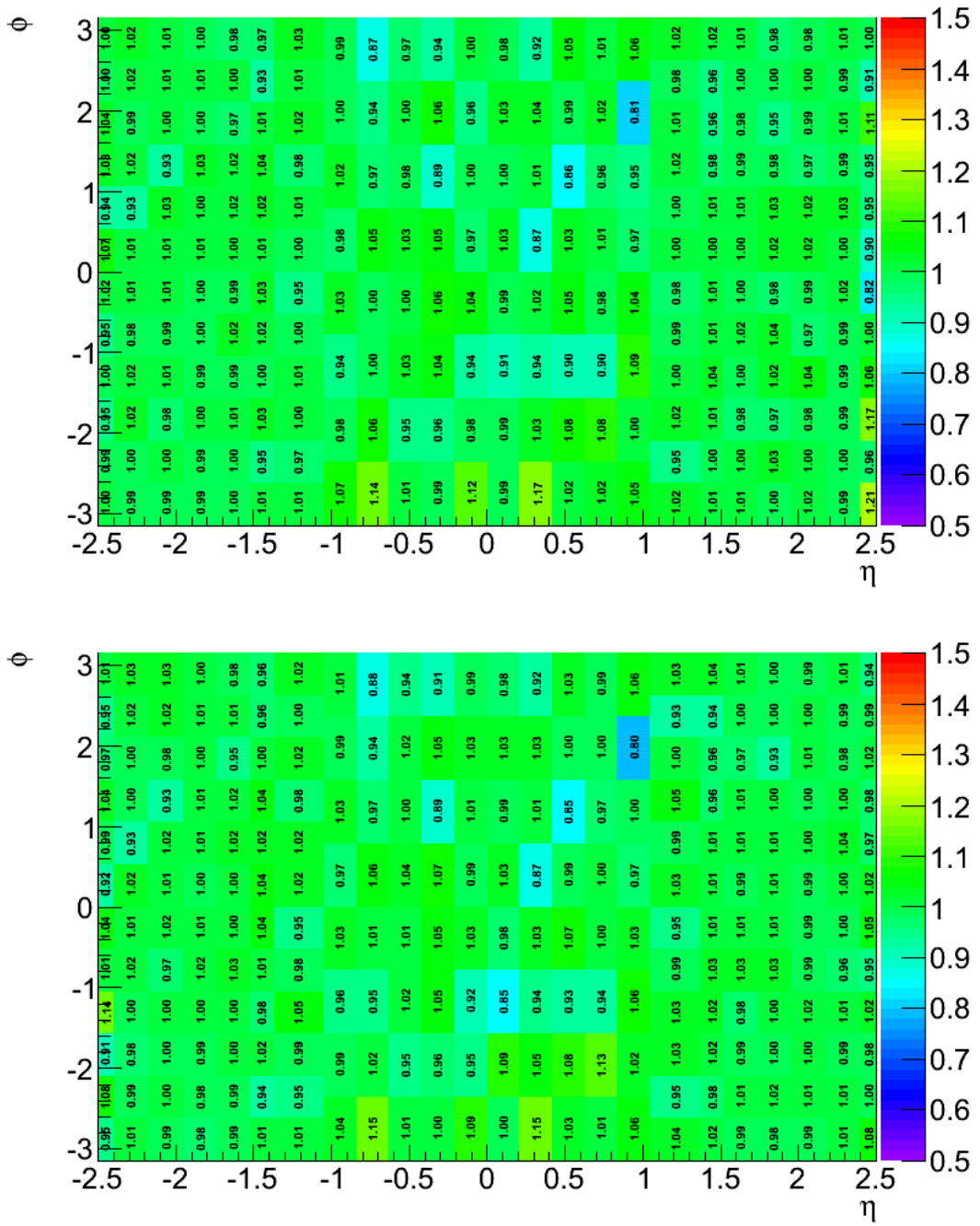

Figure A.7: Derived $\eta - \phi$ SF for periods J–M (excl. L3 and L4), given separately for positively (top) and negatively (bottom) charged muons.

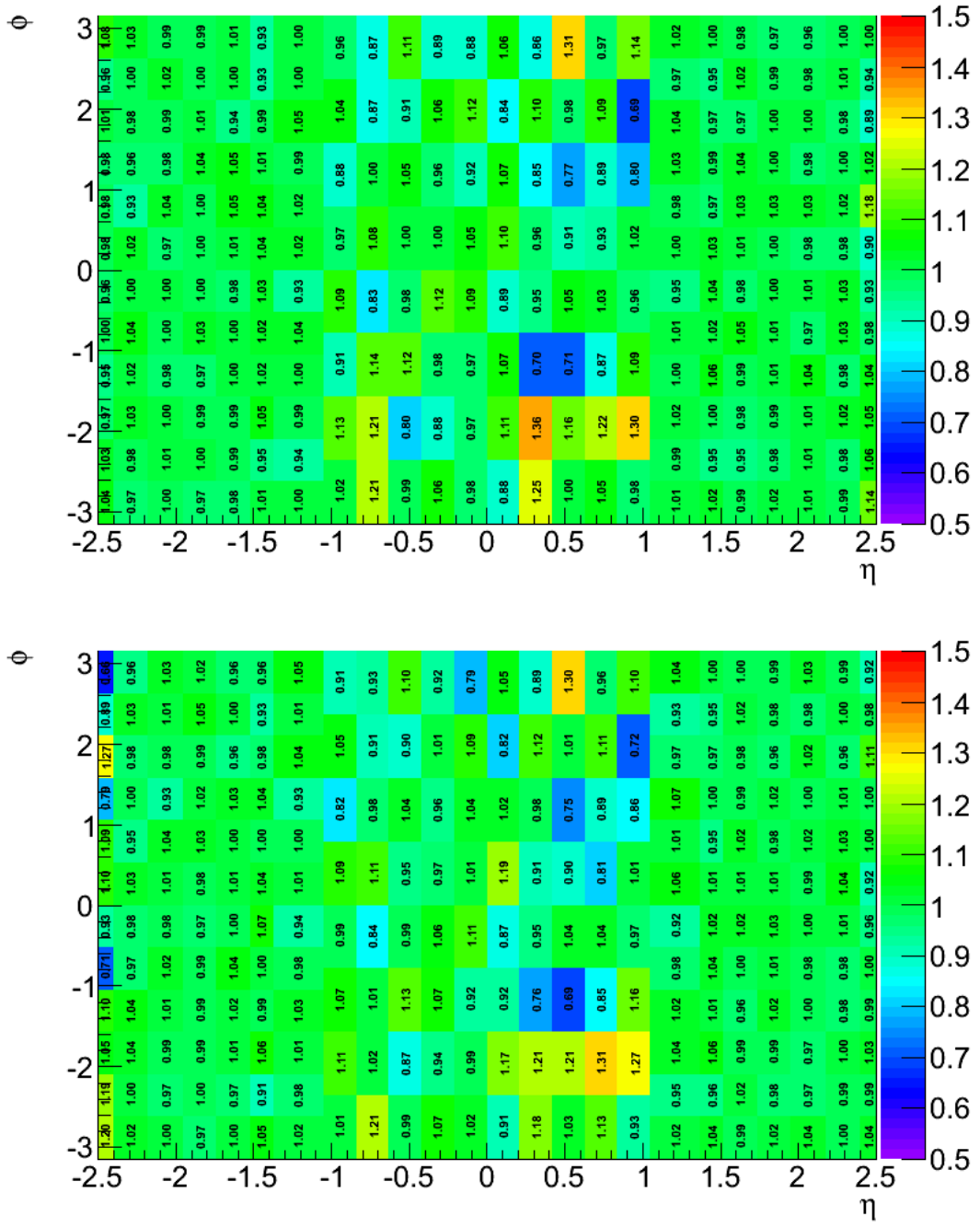

Figure A.8: Derived $\eta - \phi$ SF for periods L3 and L4, given separately for positively (top) and negatively (bottom) charged muons.

Appendix B

Nuisance Parameters

B.1 Statistical

- `norm_LH[11/12]_Top_boost` - Statistical uncertainty of top normalisation in the boosted category, as constrained by single bin top CR.
- `norm_LH[11/12]_Top_vbf` - Statistical uncertainty of top normalisation in the VBF category, as constrained by single bin top CR.
- `norm_LH[11/12]_Ztt_bv` - Statistical uncertainty of $Z \rightarrow \tau\tau$ normalisation, predominantly constrained by low score region of the BDT distribution.
- `stat_lh[11/12]_boost_bin_[n]` - Statistical uncertainty of bin n in the boosted signal region BDT distribution.
- `stat_lh[11/12]_vbf_bin_[n]` - Statistical uncertainty of bin n in the VBF signal region BDT distribution.

B.2 Systematic

- `ANA_EMB_ISOL_20[11/12]` - $Z \rightarrow \mu\mu$ selection systematic of embedding background, evaluated by varying isolation requirement of muons.
- `ANA_EMB_MFS_20[11/12]` - Embedding systematic arising from varying size of calorimeter cell subtraction.
- `ANA_LH11_Fake_boost_CLOSE` - Closure systematic uncertainty for fakes background in boosted category, taken from discrepancies observed when FF are measured in and applied to MC samples.

- `ANA_LH11_Fake_boost_RW` - Uncertainty due to variation of W+jets contribution to FF for the boosted category, determined by measuring R_W in W-rich and W-depleted CRs.
- `ANA_LH11_Fake_vbf_CLOSE` - Closure systematic uncertainty for fakes background in VBF category, taken from discrepancies observed when FF are measured in and applied to MC samples.
- `ANA_LH11_Fake_vbf_RW` - Uncertainty due to variation of W+jets contribution to FF for the VBF category, determined by measuring R_W in W-rich and W-depleted CRs.
- `ANA_LH[11/12]_Fake_boost_COMP` - Background composition uncertainty for FF in boosted category, derived by varying contribution from each fake tau background.
- `ANA_LH[11/12]_Fake_boost_STAT[n]` - Statistical uncertainty on the FF in p_T bin n for the boosted signal region, due to statistical uncertainties of the CR used to derive the FF.
- `ANA_LH[11/12]_Fake_btop_STAT[n]` - Statistical uncertainty on the FF in p_T bin n for the boosted top control region, due to statistical uncertainties of the CR used to derive the FF.
- `ANA_LH[11/12]_Fake_vbf_COMP` - Background composition uncertainty for FF in VBF category, derived by varying contribution from each fake tau background.
- `ANA_LH[11/12]_Fake_vbf_STAT[n]` - Statistical uncertainty on the FF in p_T bin n for the VBF signal region, due to statistical uncertainties of the CR used to derive the FF.
- `ANA_LH[11/12]_Fake_vtop_STAT[n]` - Statistical uncertainty on the FF in p_T bin n for the VBF top control region, due to statistical uncertainties of the CR used to derive the FF.
- `ANA_LH[11/12]_Top_boost_MCNLO` - Uncertainty due to observed differences between POWHEG and MC@NLO top samples, when extrapolating from top CR to boosted signal region.
- `ANA_LH[11/12]_Top_vbf_MCNLO` - Uncertainty due to observed differences between POWHEG and MC@NLO top samples, when extrapolating from top CR to VBF signal region.

- `ANA_LH[11/12]_Zlt_vbf_DETAJJ` - Systematic uncertainty on $|\Delta\eta_{jj}|$ reweighting applied to $Z \to \ell\ell$ ($\ell \to \tau$) events in the VBF category.
- `BTag_B[n]_20[11/12]` - Component n of b-tagging systematics for b-quark initiated jets.
- `BTag_C[n]_20[11/12]` - Component n of b-tagging systematics for c-quark initiated jets.
- `BTag_L[n]_20[11/12]` - Component n of b-tagging systematics for light-quark initiated jets.
- `EL_EFF_20[11/12]` - Systematic uncertainty on electron trigger, identification and isolation efficiencies, added in quadrature.
- `EL_RES` - Electron energy resolution systematic.
- `GenAccq2Z` - QCD scale uncertainty of Z acceptance.
- `JER` - Jet energy resolution systematic.
- `JES_20[11/12]_Eta_StatMethod` - JES uncertainty from statistical uncertainty when measuring JES for different η ranges.
- `JES_20[11/12]_Modelling1` - JES uncertainty from modelling shortcomings when measuring JES.
- `JES_20[11/12]_Statistical1` - JES uncertainty from statistical uncertainty of JES determination.
- `JES_2012_PileRho_TAU_GG` - JES uncertainty from variation with pile-up for gg initiated processes.
- `JES_2012_PileRho_TAU_QG` - JES uncertainty from variation with pile-up for qg initiated processes.
- `JES_2012_PileRho_TAU_QQ` - JES uncertainty from variation with pile-up for qq initiated processes.
- `JES_Detector1` - JES uncertainty from imperfect detector simulation when measuring JES.
- `JES_Eta_Modelling` - JES uncertainty from variation as a function of η.

- JES_Flavb - JES uncertainty from uncertainty of detector response to b quark initiated jets.
- JES_FlavComp_TAU_G - JES uncertainty from flavour composition of gluon initiated processes.
- JES_FlavComp_TAU_Q - JES uncertainty from flavour composition of quark initiated processes.
- JES_FlavResp - JES uncertainty from varying detector response to different jet flavour compositions.
- JES_Mu - JES uncertainty from variation with mean interactions per bunch crossing.
- JES_NPV - JES uncertainty from variation with number of primary vertices identified in an event.
- LUMI_20[11/12] - Systematic uncertainty on measurement of integrated luminosity of the dataset.
- MET_RESOSOFT_20[11/12] - Energy resolution uncertainty for E_T^{miss}.
- MET_SCALESOFT_20[11/12] - Energy scale uncertainty for E_T^{miss}.
- MU_EFF_20[11/12] - Systematic uncertainty on muon trigger, identification and isolation efficiencies, added in quadrature.
- MU_EFF_2012_Emb - Systematic uncertainty on muon trigger, identification and isolation efficiencies, added in quadrature, for event selection stage of embedding procedure.
- PU_RESCALE_20[11/12] - Uncertainty originating from pile-up reweighting which is applied to all simulated samples.
- TAU_EFAKE_20[11/12] - Uncertainty on $e \to \tau$ fake rates.
- TAU_ID_20[11/12] - Uncertainty on tau identification SF.
- TAU_ID_STAT_2012 - Statistical uncertainty on tau identification SF.
- TES_Fake_20[11/12] - Energy scale uncertainty of fake taus.
- TES_InSitu_2012 - TES uncertainty component arising from the in-situ tau energy correction.

- `TES_Model_2012` - TES uncertainty component due to modelling uncertainties.
- `TES_True_2011` - TES uncertainty component for true taus.
- `UE_gg` - Underlying event uncertainty for gg initiated signal processes (ggF).
- `UE_qq` - Underlying event uncertainty for qq initiated signal processes (VBF, VH).

B.3 Theoretical

- `BR_tautau` - Theoretical uncertainty on the $H \to \tau\tau$ branching ratio.
- `pdf_Higgs_gg` - PDF uncertainty for gluon initiated Higgs production processes (ggF), defined using variation of MSTW2008NLO, NNPDF and CT10 eigen-tune parameterisation.
- `pdf_Higgs_qq` - PDF uncertainty for quark initiated Higgs production processes (VBF, VH), defined using variation of MSTW2008NLO, NNPDF and CT10 eigen-tune parameterisation.
- `pdf_qq` - PDF uncertainty for quark initiated background processes modelled by MC (diboson, $Z \to \ell\ell$ ($\ell \to \tau$)).
- `QCDscale_ggH_m12` - QCD scale uncertainty for ggH + $\geq$ 2 jets events.
- `QCDscale_ggH_m23` - QCD scale uncertainty for ggH + $\geq$ 3 jets events.
- `QCDscale_ggH_ptH_m01` - QCD scale uncertainty for ggH + $\geq$ 1 jets events with $p\_T^H > 100$ GeV.
- `QCDscale_qqH` - QCD scale uncertainty for VBF events.

Bibliography

[1] K. Olive, K. Agashe, C. Amsler, et al. Review of Particle Physics. *Chinese Physics C*, **38**(9):090001, (2014).

[2] F. Halzen and A. D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, Inc., (1984).

[3] Q. R. Ahmad, R. C. Allen, T. C. Andersen, et al. Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory. *Physical Review Letters*, **87**:071301, (2001).

[4] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Physical Review Letters*, **13**:321–323, (1964).

[5] P. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Physical Review Letters*, **13**:508–509, (1964).

[6] G. Guralnik, C. Hagen, and T. Kibble. Global Conservation Laws and Massless Particles. *Physical Review Letters*, **13**:585–587, (1964).

[7] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, **716**:1–29, (2012).

[8] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, **716**:30–61, (2012).

[9] The ATLAS Collaboration. Measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector. *Physical Review D*, **90**:052004, (2014).

[10] The CMS Collaboration. Precise determination of the mass of the Higgs boson and studies of the compatibility of its couplings with the standard model. In *CMS-PAS-HIG-14-009*, (2014).

[11] The ATLAS Collaboration. Evidence for the spin-0 nature of the Higgs boson using ATLAS data. *Physics Letters B*, **726**:120–144, (2013).

[12] The CMS Collaboration. Study of the mass and spin-parity of the Higgs boson candidate via its decays to Z boson pairs. *Physical Review Letters*, **110**:081803, (2013).

[13] D. Griffiths. *Introduction to Elementary Particles*. John Wiley & Sons, Inc., (1987).

[14] C. Yang and R. Mills. Conservation of Isotopic Spin and Isotopic Gauge Invariance. *Physical Review*, **96**:191–195, (1954).

[15] R. L. Liboff. *Introductory Quantum Mechanics*. Addison-Wesley, (2002).

[16] H. Georgi. *Lie Algebras in Particle Physics*. Westview Press, second edition, (1999).

[17] J. Goldstone, A. Salam, and S. Weinberg. Broken Symmetries. *Physical Review*, **127**:965–970, (1962).

[18] I. J. R. Aitchison and A. J. G. Hey. *Gauge Theories in Particle Physics, Volume 2: Non-Abelian Gauge Theories: QCD and the Electroweak Theory*. Taylor & Francis Group, fourth edition, (2013).

[19] The LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. *CERN-2011-002*, (2011).

[20] `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections`.

[21] D. Binosi and L. Theußl. JaxoDraw: A graphical user interface for drawing Feynman diagrams. *Computer Physics Communications*, **161**:76–86, (2004).

[22] `http://home.web.cern.ch/about/accelerators/linear-accelerator-2`.

[23] `http://home.web.cern.ch/about/accelerators/proton-synchrotron-booster`.

[24] `http://home.web.cern.ch/about/accelerators/proton-synchrotron`.

[25] `http://home.web.cern.ch/about/accelerators/super-proton-synchrotron`.

[26] `http://home.web.cern.ch/about/accelerators/large-hadron-collider`.

[27] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[28] The LHCb Collaboration. The LHCb Detector at the LHC. *Journal of Instrumentation*, **3**:S08005, (2008).

[29] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, **3**:S08003, (2008).

[30] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, **3**:S08004, (2008).

[31] The ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, **3**:S08002, (2008).

[32] http://press.web.cern.ch/press-releases/2008/10/cern-releases-analysis-lhc-incident.

[33] M. Aleksa, F. Bergsma, P. A. Giudici, et al. Measurement of the ATLAS Solenoid Magnetic field. *Journal of Instrumentation*, **3**:P04003, (2008).

[34] The ATLAS Collaboration. ATLAS Muon Spectrometer Technical Design Report. *CERN-LHCC-97-022*, (1997).

[35] The ATLAS Collaboration. The ATLAS Inner Detector commissioning and calibration. *European Physical Journal C*, **70**(3):787–821, (2010).

[36] The ATLAS Collaboration. Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector. *ATLAS-CONF-2011-063*, (2011).

[37] The ATLAS Collaboration. Performance of the ATLAS muon trigger in 2011. *ATLAS-CONF-2012-099*, (2012).

[38] The ATLAS Collaboration. Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV. *European Physical Journal C*, **75**:120, (2015).

[39] The ATLAS Collaboration. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data. *European Physical Journal C*, **74**:3130, (2014).

[40] The ATLAS Collaboration. Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data. *ATLAS-CONF-2014-032*, (2014).

[41] W. Lampl, S. Laplace, D. Lelas, et al. Calorimeter Clustering Algorithms: Description and Performance. *ATL-LARG-PUB-2008-002*, (2008).

[42] The ATLAS Collaboration. Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data. *European Physical Journal C*, **74**:2941, (2014).

[43] The ATLAS Collaboration. Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data. *European Physical Journal C*, **74**:3071, (2014).

[44] M. Cacciari, G. P. Salam, and G. Soyez. The anti-k_T jet clustering algorithm. *Journal of High Energy Physics*, **04**:063, (2008).

[45] S. D. Ellis and D. E. Soper. Successive Combination Jet Algorithm for Hadron Collisions. *Physical Review D*, **48**:3160–3166, (1993).

[46] The ATLAS Collaboration. Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment. *ATLAS-CONF-2014-004*, (2014).

[47] The ATLAS Collaboration. Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8$ TeV. *ATLAS-CONF-2015-017*, (2015).

[48] The ATLAS Collaboration. Tagging and suppression of pileup jets with the ATLAS detector. *ATLAS-CONF-2014-018*, (2014).

[49] The ATLAS Collaboration. Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s} = 8$ TeV. *European Physical Journal C*, **75**:303, (2015).

[50] The ATLAS Collaboration. Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at $\sqrt{s} = 8$ TeV. *ATLAS-CONF-2013-082*, (2013).

[51] A. Elagin, P. Murat, A. Pranko, and A. Safonov. A New Mass Reconstruction Technique for Resonances Decaying to $\tau\tau$. *Nuclear Instruments and Methods in Physics Research A*, **654**:481–489, (2011).

[52] The ATLAS Collaboration. Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector. *Journal of High Energy Physics*, **04**:117, (2015).

[53] P. Nason. A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms. *Journal of High Energy Physics*, **11**:040, (2004).

[54] S. Frixione, P. Nason, and C. Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *Journal of High Energy Physics*, **11**:070, (2007).

[55] S. Alioli, P. Nason, C. Oleari, and E. Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *Journal of High Energy Physics*, **06**:043, (2010).

[56] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini. Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM. *Journal of High Energy Physics*, **02**:088, (2012).

[57] T. Sjöstrand, S. Mrenna, and P. Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, **178**:852–867, (2008).

[58] H. L. Lai, M. Guzzi, J. Huston, et al. New parton distributions for collider physics. *Physical Review D*, **82**:074024, (2010).

[59] A. Djouadi, M. Spira, and P. Zerwas. Production of Higgs bosons in proton colliders. QCD corrections. *Physics Letters B*, **264**:440–446, (1991).

[60] S. Dawson. Radiative corrections to Higgs boson production. *Nuclear Physics B*, **359**:283–300, (1991).

[61] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas. Higgs boson production at the LHC. *Nuclear Physics B*, **453**:17–82, (1995).

[62] R. V. Harlander and W. B. Kilgore. Next-to-Next-to-Leading Order Higgs Production at Hadron Colliders. *Physical Review Letters*, **88**:201801, (2002).

[63] C. Anastasiou and K. Melnikov. Higgs boson production at hadron colliders in NNLO QCD. *Nuclear Physics B*, **646**:220–256, (2002).

[64] V. Ravindran, J. Smith, and W. L. van Neerven. NNLO corrections to the total cross section for Higgs boson production in hadron-hadron collisions. *Nuclear Physics B*, **665**:325–366, (2003).

[65] S. Catani, D. de Florian, M. Grazzini, and P. Nason. Soft-gluon resummation for Higgs boson production at hadron colliders. *Journal of High Energy Physics*, **07**:028, (2003).

[66] U. Aglietti, R. Bonciani, G. Degrassi, and A. Vicini. Two-loop light fermion contribution to Higgs production and decays. *Physics Letters B*, **595**:432–441, (2004).

[67] S. Actis, G. Passarino, C. Sturm, and S. Uccirati. NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders. *Physics Letters B*, **670**:12–17, (2008).

[68] M. Grazzini and H. Sargsyan. Heavy-quark mass effects in Higgs boson production at the LHC. *Journal of High Energy Physics*, **09**:129, (2013).

[69] D. De Florian, G. Ferrera, M. Grazzini, and D. Tommasini. Higgs boson production at the LHC: transverse momentum resummation effects in the $H \to \gamma\gamma$, $H \to WW \to l\nu l\nu$ and $H \to ZZ \to 4l$ decay modes. *Journal of High Energy Physics*, **06**:132, (2012).

[70] J. M. Campbell, R. K. Ellis, and G. Zanderighi. Next-to-Leading order Higgs + 2 jet production via gluon fusion. *Journal of High Energy Physics*, **10**:028, (2006).

[71] A. Banfi, P. F. Monni, G. P. Salam, and G. Zanderighi. Higgs and Z-boson production with a jet veto. *Physical Review Letters*, **109**:202001, (2012).

[72] A. Banfi, G. P. Salam, and G. Zanderighi. NLL+NNLO predictions for jet-veto efficiencies in Higgs-boson and Drell-Yan production. *Journal of High Energy Physics*, **06**:159, (2012).

[73] A. Banfi, P. F. Monni, and G. Zanderighi. Quark masses in Higgs production with a jet veto. *Journal of High Energy Physics*, **01**:097, (2014).

[74] M. Zaro, P. Bolzoni, F. Maltoni, and S.-O. Moch. Higgs production via vector-boson fusion at NNLO in QCD. *Physical Review Letters*, **105**:011801, (2010).

[75] M. Ciccolini, A. Denner, and S. Dittmaier. Strong and electroweak corrections to the production of Higgs+2jets via weak interactions at the LHC. *Physical Review Letters*, **99**:161803, (2007).

[76] M. Ciccolini, A. Denner, and S. Dittmaier. Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC. *Physical Review D*, **77**:013002, (2008).

[77] K. Arnold, M. Bähr, G. Bozzi, et al. VBFNLO: A parton level Monte Carlo for processes with electroweak bosons. *Computer Physics Communications*, **180**:1661–1670, (2009).

[78] The LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 2. Differential Distributions. *CERN-2012-002*, (2012).

[79] J. Pumplin, D. R. Stump, J. Huston, et al. New Generation of Parton Distributions with Uncertainties from Glaobal QCD Analysis. *Journal of High Energy Physics*, **07**:012, (2002).

[80] O. Brein, A. Djouadi, and R. Harlander. NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders. *Physics Letters B*, **579**:149–156, (2004).

[81] M. L. Ciccolini, S. Dittmaier, and M. Krämer. Electroweak Radiative Corrections to Associated WH and ZH Production at Hadron Colliders. *Physical Review D*, **68**:073003, (2003).

[82] The ATLAS Collaboration. Summary of ATLAS Pythia 8 tunes. *ATL-PHYS-PUB-2012-003*, (2012).

[83] The ATLAS Collaboration. The ATLAS Simulation Infrastructure. *European Physical Journal C*, **70**(3):823–874, (2010).

[84] GEANT4 Collaboration. GEANT4 - A Simulation toolkit. *Nuclear Instruments and Methods in Physics Research A*, **506**:250–303, (2003).

[85] The ATLAS Collaboration. Measurement of the inclusive $W^{\pm}$ and Z/γ^{*} cross sections in the e and μ decay channels in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Physical Review D*, **85**:072004, (2012).

[86] S. Jadach, Z. Was, R. Decker, and J. Kühn. The τ decay library TAUOLA, version 2.4. *Computer Physics Communications*, **76**:361–380, (1993).

[87] Z. Czyczula, T. Przedzinski, and Z. Was. TauSpinner program for studies on spin effect in tau production at the LHC. *European Physical Journal C*, **72**:1988, (2012).

[88] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *Journal of High Energy Physics*, **07**:001, (2003).

[89] G. Corcella, I. G. Knowles, G. Marchesini, et al. HERWIG 6.5 Release Note, (2002).

[90] B. P. Kersevan and E. Richter-Was. The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1. *Computer Physics Communications*, **184**:919–985, (2013).

[91] T. Sjostrand, S. Mrenna, and P. Skands. PYTHIA 6.4 Physics and Manual. *Journal of High Energy Physics*, **05**:026, (2006).

[92] T. Binoth, M. Ciccolini, N. Kauer, and M. Krämer. Gluon-induced W-boson pair production at the LHC. *Journal of High Energy Physics*, **12**:046, (2006).

[93] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen. *Classification and Regression Trees*. CRC Press, (1984).

[94] C. Gini. *Variabilità e mutabilità, Reprinted in Memorie di Metodologia Statistica*. Rome: Libreria Eredi Virgilio Veschi, (1955).

[95] J. H. Friedman. Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*, **29**(5):1189–1232, (2001).

[96] A. Hoecker, P. Speckmayer, J. Stelzer, et al. TMVA - Toolkit for Multivariate Data Analysis. *Proceedings of Science (Advanced Computing and Analysis Techniques)*, **040**, (2007).

[97] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke. HistFactory: A tool for creating statistical models for use with RooFit and RooStats. *CERN-OPEN-2012-016*, (2012).

[98] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *European Physical Journal C*, **71**:1554, (2011).

[99] R. Brun and F. Rademakers. ROOT - An Object Oriented Data Analysis Framework. *Nuclear Instruments and Methods in Physics Research A*, **389**:81–86, (1997).

[100] The ATLAS Collaboration. Improved luminosity determination in *pp* collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC. *European Physical Journal C*, **73**:2518, (2013).

[101] The ATLAS Collaboration. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. *European Physical Journal C*, **73**:2304, (2013).

[102] The ATLAS Collaboration. Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at $\sqrt{s} = 7$ TeV in 2011. *ATLAS-CONF-2012-048*, (2012).

[103] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *European Physical Journal C*, **63**(2):189–285, (2009).

[104] R. D. Ball, V. Bertone, F. Cerutti, et al. Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology. *Nuclear Physics B*, **849**:296–363, (2011).

[105] S. Frixione and B. R. Webber. Matching NLO QCD Computations and Parton Shower Simulations. *Journal of High Energy Physics*, **06**:029, (2002).

[106] J. Alwall, R. Frederix, S. Frixione, et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, **07**:079, (2014).